

Copyright
by
Jay Brady Fletcher
2013

**THE DISSERTATION COMMITTEE FOR JAY BRADY FLETCHER
CERTIFIES THAT THIS IS THE APPROVED VERSION OF THE FOLLOWING
DISSERTATION:**

**CONTROL AND IMPLEMENTATION OF INTEGRATED
VOLTAGE REGULATORS**

Committee:

Earl Swartzlander, Jr., Supervisor

Mircea Driga

Mack Grady

Shawn Searles

Nur Touba

**CONTROL AND IMPLEMENTATION OF INTEGRATED
VOLTAGE REGULATORS**

by

JAY BRADY FLETCHER, B.S.; M.S.E.

DISSERTATION

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN

DECEMBER 2013

Acknowledgements

I appreciate my advisor, Earl Swartzlander, Jr for his enduring support and words of encouragement. Additionally, I would like to thank my friends and family for helping me through this endeavor and pushing me along. Finally, I greatly appreciate the technical leaders and managers at AMD in Austin for financing the bulk of this graduate work and allowing me to work on it.

CONTROL AND IMPLEMENTATION OF INTEGRATED VOLTAGE REGULATORS

Jay Brady Fletcher, Ph.D.

The University of Texas at Austin, 2013

Supervisor: Earl Swartzlander, Jr.

This dissertation describes the development of voltage regulators for the purpose of power reduction and further scaling in highly integrated system-on-chip products. Emphasis is placed on the architecture and implementation of integrated voltage regulation using commercially available components, standard CMOS technology, and a practical controller. The research spans the fundamental elements, architectural aspects, and detailed analog integrated circuit design.

TABLE OF CONTENTS

TABLE OF CONTENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
1 Introduction	1
1.1 Parallelization and Integration Trends	1
1.2 The Importance of Power	4
1.3 Power Granularity	6
1.4 Static and Dynamic Currents	7
1.5 Pin and Package Resources	7
1.6 Technology vs. Analog Circuits	8
1.7 Fully Integrated Voltage Regulation	10
1.8 Summary	10
2 Background	12
2.1 Linear Regulation	12
2.2 1897: Nikola Tesla's Currents of Ordinary Character	14
2.3 Sixteen-Phase Integrated Solution	16
2.4 AMD 32 nm SOI Capacitive Converter	17
2.5 IBM 45 nm Deep Trench Capacitor Converter	18
2.6 Conventional Voltage Regulator Down (VRD) Solution	19
2.7 Intel Thin Film Ferrite Multiple Chip Module	23
2.8 Intel Package Inductor MCM	24
2.9 Summary	25

3	Switched-Capacitor Solutions	27
3.1	Generalized IVR Control System	27
3.2	Converter Modeling	29
3.2.1	Manual Solver	29
3.2.2	MATLAB and Circuit Cosimulator	32
3.3	Switched Capacitor Usage	36
4	Buck Elements	38
4.1	Inductor	38
4.1.1	Integrated Magnetic Power Inductors	39
4.1.2	Discrete Inductors	41
4.2	Capacitor	49
4.2.1	Integrated Capacitor Options	50
4.3	Power Switches	52
4.3.1	Performance Metrics	53
4.3.2	Single stacked Power FETs	58
4.3.3	Double Stacked Power FETs	60
4.3.4	Triple Stacked Power FETs	63
4.3.5	Summary of Stacking Capability	66
4.4	Summary	68
5	Architecture	69
5.1	IVR Practicality	69
5.2	Load Definition	70
5.2.1	Physical Constraints	70
5.2.2	Electrical Constraints	75
5.3	Phase-Centric Designs	79
5.3.1	Coilcraft 0908SQ-27N_L Phase	80

5.3.2 Coilcraft 0807SQ-11N_L Solution	84
5.4 Voltage Droop.....	85
5.5 Load Line and Automatic Voltage Positioning	91
5.6 Efficiency	95
5.7 Summary of Architecture.....	100
6 Circuit Design	101
6.1 Controller Strategy.....	102
6.1.1 Operational Modes.....	106
6.2 Block Circuits	110
6.2.1 Power FETs	110
6.2.2 Level Shifter and Non-overlapping Clock Generator (LSNOC)	112
6.2.3 Current Sensor	115
6.2.4 Current Loop Stabilizer	118
6.2.5 Current Comparison Circuit.....	119
6.2.6 Voltage Error Amplifier.....	122
6.3 Input Filtering	123
6.4 Circuit Summary	124
7 Conclusions.....	125
7.1 Advantages.....	126
7.2 Disadvantages	127
7.3 Future Work	128
7.3.1 Fully Integrated Possibilities	128
7.3.2 CMOS and Supply Voltage Departure	128
7.3.3 Architectural Features for Low Power.....	129
BIBLIOGRAPHY	130

LIST OF TABLES

Table 1: Common classes of IP	5
Table 2: Conventional regulator design specifications	22
Table 3: Summary of prior power conversion solutions	26
Table 4: Power FET characterization data from ASU 45 nm PTM model	56
Table 5: Stacked power FET performance summary	67
Table 6: Tuolumne physical specifications	74
Table 7: Tuolumne CPU electrical load constraints	77
Table 8: Tuolumne GPU electrical load constraints	78
Table 9: Comparison of conventional and integrated voltage regulators.	125

LIST OF FIGURES

Figure 1: An AMD 32 nm quad-core accelerated processing unit (APU) with integrated graphics processor	2
Figure 2: An Intel Architecture single-chip cloud computer with 48 core-pairs and area of 567 mm ² [25].	2
Figure 3: Power hierarchy of a legacy system.	4
Figure 4: Diagram of fully integrated power delivery	8
Figure 5: Intel Fin-FET transistors [14]	9
Figure 6: Power efficiency versus output current of a linear regulator with $V_{DD} =$ 1.5 V, $V_{load} = 1.2$ V, and $I_q = 100$ μ A	13
Figure 7: Nikola Tesla's “Apparatus for Producing Currents of High Frequency” from 1897 [57]	15
Figure 8: Depiction of the trade-off between efficiency and power density in switched capacitor regulators (from [33])	18
Figure 9: IBM deep trench capacitor converter (from [9])	19
Figure 10: Intel LGA1366 recommended motherboard power routing (from [28])	21
Figure 11: Intel thin film Ni ₈₀ Fe ₂₀ magnetic solution (from [17])	24
Figure 12: Intel package inductor solutions (from [53])	24
Figure 13: Switched-capacitor integrated regulation system.....	28
Figure 14: Single-cell circuit schematic [33]	29
Figure 15: Comparison of single-cell DC simulations in MATLAB and Cadence Spectre.....	31
Figure 16: Single-cell converter mathematical model clocked in 2:1 conversion mode	32

Figure 17: Circuit and behavioral element interaction in the automatic MATLAB hybrid model	33
Figure 18: A simple closed loop capacitive converter model	34
Figure 19: Output voltage versus time of the MATLAB model in Figure 18	35
Figure 20: Input listing to MATLAB cosimulator for switched-capacitor development	36
Figure 21: Vishay inductors for high current motherboard power supplies	39
Figure 22: Thin-film strip-style power inductor (from [38])	40
Figure 23: Thin-film yolk-style power inductor (from [57])	40
Figure 24: Wire wound ferrite core inductor (from [39])	42
Figure 25: Current density of 0603 and 0402 ferrite core wire wound inductors ..	43
Figure 26: Wire-wound air core inductors (from [12])	44
Figure 27: DCR versus inductance for a series of air core inductors	44
Figure 28: Thin-film ferrite discrete inductor (from [50])	45
Figure 29: Current density versus inductance for three discrete inductor structures ..	46
Figure 30: Current density versus DC resistance of inductor types	47
Figure 31: Conductance versus inductance of inductor types	48
Figure 32: Conductance versus f_{sw} of inductor types	49
Figure 33: Tiny built-in capacitors (from [39])	50
Figure 34: Intel 22 nm process MIM capacitor between M8 and M9 (from [5]) ..	51
Figure 35: Full metal stack of Intel 22 nm Tri-Gate (from [5])	51
Figure 36: Power FET drain-source conductance versus Q_{gate}	55
Figure 37: Schematic of single stacked power FETs	58
Figure 38: Switching waveform of single stacked power FET topology	59
Figure 39: Schematic of double stacked power FETs.	61

Figure 40: Switching waveform of double stack power FETs.	62
Figure 41: Schematic of triple stacked power FETs for higher input supply.	64
Figure 42: Switching waveform of the triple stack power FET topology	65
Figure 43: Close-packed hexagonal bump spacing.	71
Figure 44: Simple grid structure	72
Figure 45: Scale drawing of the Tuolumne high performance server chip.....	75
Figure 46: Scale drawing of 40 Coilcraft 0908SQ-27N_L 4.4 A inductor placements	81
Figure 47: Scale drawing of Coilcraft 0807SQ-11N_L 2.7 A inductor placements	84
Figure 48: Schematic of regulator with load step	86
Figure 49: Reduced model for derivation of maximum V_{droop} as a result of instantaneous current step	87
Figure 50: Voltage droop according to inverse Laplace model.....	89
Figure 51: Positive and negative current steps with and without AVP.	93
Figure 52: Graph of a 2.0 m Ω load line for a high performance microprocessor .	94
Figure 53: Breakdown of power loss in 27 nH phase	95
Figure 54: Power efficiency versus output current	96
Figure 55: Semi-log plot of power efficiency versus current	97
Figure 56: Improvement in efficiency through phase shedding	99
Figure 57: The complete IVR system.	102
Figure 58: Voltage mode controlled buck topology.	103
Figure 59: Current mode controlled buck topology.....	105
Figure 60: CCM current waveform.....	107
Figure 61: Negative CCM current waveform	108
Figure 62: DCM current waveform	109

Figure 63: Power FETs within current mode loop.....	111
Figure 64: Schematic of power FET MCP and MCN gate-bias amplifier	112
Figure 65: LSNOG within current mode loop	113
Figure 66: Non-overlap logic	113
Figure 67: Level shift and non-overlapping tapered buffer stage circuit.....	114
Figure 68: Current sensor within current mode loop	115
Figure 69: Current sensor topology	116
Figure 70: Schematic of current sensor circuit (track and hold not shown)	117
Figure 71: Current loop stabilizer within current mode loop	118
Figure 72: Schematic of current stabilizer	119
Figure 73: Current loop comparator within current mode loop	120
Figure 74: Schematic of current comparator circuit	121
Figure 75: Voltage error amplifier within current mode loop	122
Figure 76: Schematic of voltage error amplifier	123

1 INTRODUCTION

This work develops an integrated voltage regulation system for highly integrated monolithic systems. Such devices need the capability to deliver fine-grained power to IP subsystems without adversely adding to the total system cost or power consumption. Inductive and capacitive converters have been proposed with a focus on the energy storage mechanism and topologies.

This chapter describes the integration trend and ensuing problem of powering a highly integrated system on a chip (SOC) efficiently. Chapter 2 describes the history of integrated power solutions and linking the technology beyond a century ago. Chapter 3 discusses the use of switched-capacitor power supplies and their place in the solution space. Chapters 4, 5, and 6 develop a practical integrated power delivery solution by walking through the main elements, high-level system design, and low-level CMOS circuit design respectively. Chapter 7 discusses results, conclusions, and the potential for future work.

1.1 Parallelization and Integration Trends

The trend of parallel cores on a single monolithic die continues to grow along with increasing integration onto the SOC. A graphics unit, four cores, the northbridge, and a host of mixed signal circuits comprise AMD's 32nm Opteron, depicted in Figure 1. Intel released a 6-core part with an integrated memory controller named Westmere [30]. Additionally, Intel produced a part with 24 core-pairs, or 48 discrete CPU cores, on an enormous 567 mm² die [25].

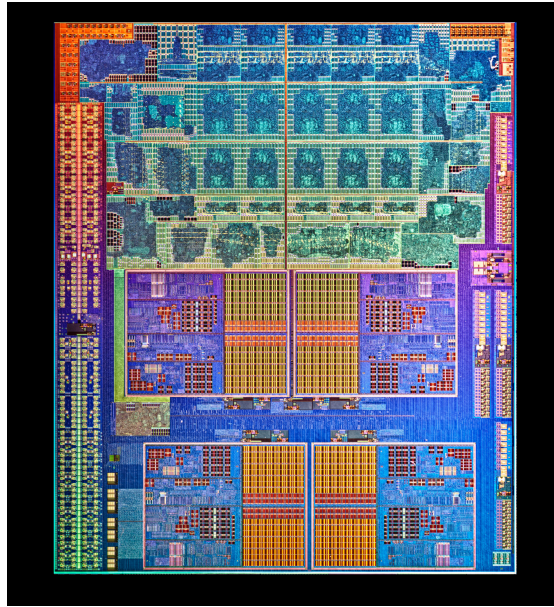


Figure 1: An AMD 32 nm quad-core accelerated processing unit (APU) with integrated graphics processor



Figure 2: An Intel Architecture single-chip cloud computer with 48 core-pairs and area of 567 mm^2 [25].

Power delivery is further complicated as PHYs/IOs, media encoders, security processors, microcontrollers, and mixed signal intellectual property (IP) are integrated onto the die. Each desires a separate power source for their respective purpose to minimize power consumption. Cache arrays, or RAMs, also prefer a separate power supply from their respective functional units. Individual power supplies allow these units to operate at their optimal voltage and frequency combination and to power down completely when not in use, lowering the total power consumption of the chip.

Mixed signal subsystems and PHYs possess constraints such as voltage, frequency, headroom, noise, or signaling levels. Supplying each major block with its own optimal power supply on the motherboard is cost and area prohibitive. Parallelization and further integration will continue in the future.

Splitting the power supply of logic and memory may result in lower power. The two types of circuits have different operational requirements as well as usage patterns. Leakage currents dominate memory arrays' power consumption while dynamic power dominates logical cores' power consumption. Splitting the logic and memory supplies may allow for a more optimal solution from a power perspective while still meeting the design targets.

All of these issues push the trend towards more voltage islands on the die. Physically separating the supplies limits their respective conductivity. Figure 3 depicts the power hierarchy of a system today.

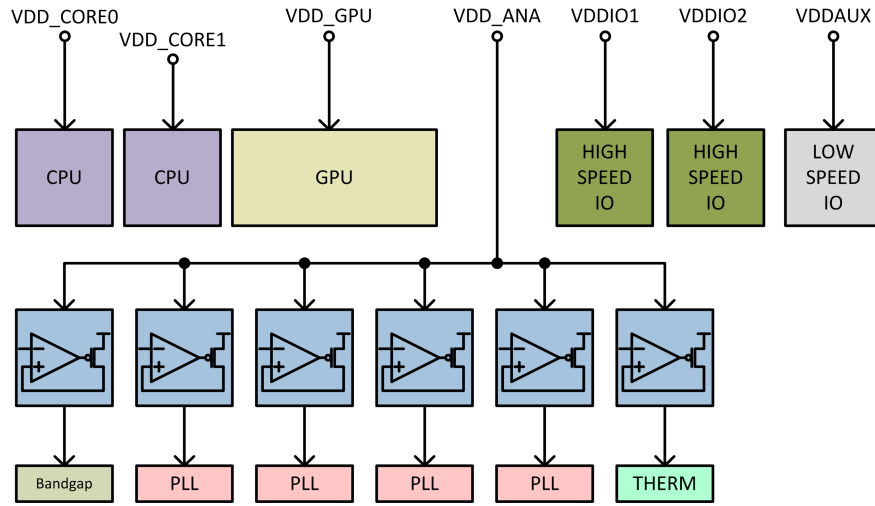


Figure 3: Power hierarchy of a legacy system.

1.2 The Importance of Power

Reducing the power consumption is the ultimate priority of system design today. Ultra-light handheld devices with brilliant displays and extensive feature lists have driven plug-in desktop computers nearly out of the market. Battery improvements proceed slowly while semiconductor technology marches forward exponentially. This has resulted in a growing dependence on battery-powered devices with advanced power management technologies. There are now only two major market segments, mobile and server, where there used to be three. Desktop computer revenue has shrunk as a result of increased performance of smaller computers.

Cost effective thermal design power (TDP) has reached a point wherein supplying more power to the processor is not a viable option.

A shift to mobile computing has also caused a sharp increase in server-side processing. Thus, the need for server-side computing solutions in “the cloud.” Amazon, Microsoft, Google, and others constructed massive compute warehouses during 2005-2010 in locations where power costs are lowest. Power in server segments proves to be

important as well as in mobile. Many features that make mobile processors power efficient are making their way into the server feature lists.

Product die areas have remained constant over many generations, 100-200 mm², driven primarily by defect densities. As technologies shrink, architects pull more features off of the PCB and into a monolithic solution.

Increased integration poses many benefits. These include lower power consumption, lower system cost, and smaller form factor. I/O drivers between parts are eliminated. Supporting circuits such as thermal monitors, ESD, PLLs, JTAG, and BIST controllers can be shared. Manufacturing and test is consolidated to some degree, albeit at a cost in complexity. Testability is challenged as integration continues to grow.

The direction of increased integration is clear. Pack as many features onto a monolithic die as the yield allows. The following is a list of IP classes that may be integrated into an SOC with examples of each.

CPU Logic	X86 Core, Northbridge, RISC Core, GPU
Memory Array	L2 Cache, DRAM, Register File
Analog	PLL, bandgap, sensors, ADC
PHY	SERDES, DRAM interface, HyperTransport

Table 1: Common classes of IP

Today, all of these functions succumb to technology optimized for digital functions. That may change in the future, as pure analog and I/O functions generally do not shrink with technology. Digitally assisted analog and I/O functions benefit from denser digital offerings.

The voltage supply for a specific IP, such as a CPU core, may need to be increased or decreased dramatically depending on the function that is being performed. In large microprocessors this is accomplished by driving a code from the processor die to a switching regulator on the motherboard, which in turn adjusts the voltage supply going back to the processor after a delay. The code may tell the regulator to disable the supply completely.

Current profiles further power distribution complexity. Digital supplies prove to be the most problematic. Leakage current, or sub-threshold conduction currents, rival or even surpass dynamic current in fast corners yet disappear at slow corners completely. Digital power supplies must be efficient and stable across a wide current output range accordingly. Digital blocks' low power states contribute significantly to battery life, which in turn requires the power supply to be efficient in low current states. Additionally, digital supplies contribute more switching noise onto the supply rails and demand significantly higher currents than their analog counterparts.

PHYs or I/O IPs generally demonstrate characteristics of both analog and digital supplies. They often incorporate analog and digital functionality and consume moderate currents. The voltage supplies do not easily scale during low power states as all digital blocks do.

1.3 Power Granularity

Clearly, a voltage source that ramps up or down quickly can save power over a slower one. With big and slow switching regulators off die the granularity of power is limited. Some people alleviate the issue with power gating. Coupling the voltage supply

closely to the destination increases the power granularity, thereby lowering the power consumption even further.

1.4 Static and Dynamic Currents

Static power may overcome dynamic power in some cases, making it difficult to bracket power consumption across corners for a given application.

1.5 Pin and Package Resources

The latest Intel LGA socket connects to the motherboard through an astounding 2,011 pins with power and ground dominating the pin count. Each power supply on the die requires significant package metal to carry current in and out of the die. These resources add cost to the overall system in terms of pin count and package complexity.

Individual supplies and signals share the total metal in the path from system board to the die. The decreased power path conductivity contributes to an overall loss in the power supply conductance at the destination circuits. DC voltage droop results in lower efficiency. Bandwidth limitations of the external power supply cause a sharp increase in power supply impedance at high frequencies. Local decoupling capacitors sprinkled throughout the die attempt to reduce the supply impedance at high frequencies at the expense of area and total cost.

Sharing the metal resources across functional units can be accomplished by pushing the final voltage supplies on die as shown in Figure 4. This architecture shares pin and package metallization among the downstream supplies rather than dividing it. Furthermore, by drawing power from a higher supply, a transformation in the required current is achieved, thereby reducing the necessary input current. This allows a reduction

in the total number of power supply pins on the package and reduces the cost of the package.

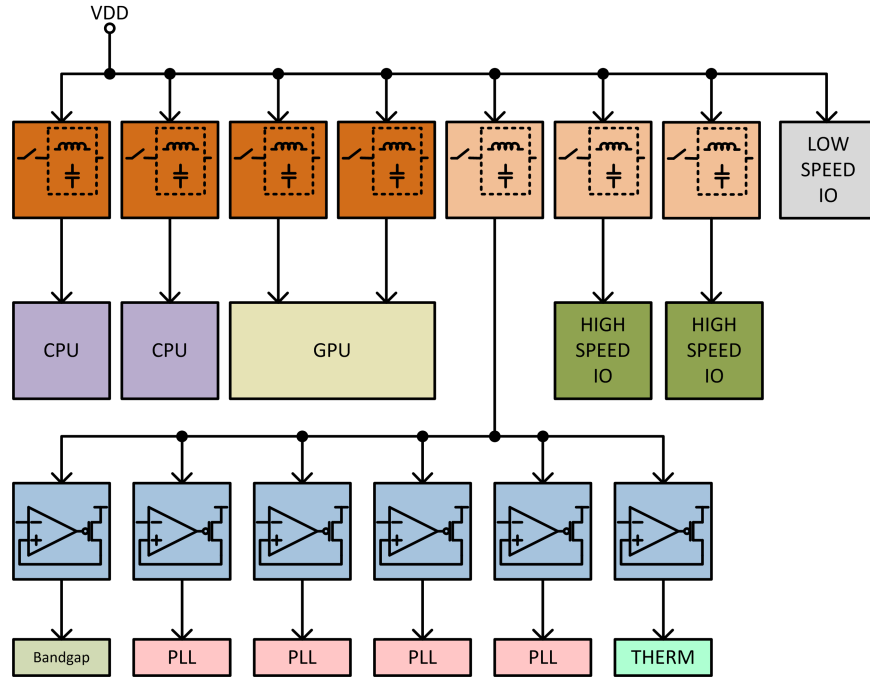


Figure 4: Diagram of fully integrated power delivery

1.6 Technology vs. Analog Circuits

The progression of CMOS technology limits the capability of analog circuits. The decreased oxide thickness increases gate leakage. Shorter channel lengths decrease r_{out} while g_m is already low, making the characteristic gain of the devices low.

An extremely wide gap exists in transistor performance across process, voltage, and temperature. This gap spread may result in overly complex or risky circuits for relatively simple applications to overcome a wide operating range.

Non-transistor elements such as resistors (poly, well, or floating), diodes, triple wells, and capacitors all require additional development and process steps. The

development of these elements in a process comes as an afterthought to digital transistor performance, even though analog designers rely on them.

Fin-FET technology will likely dominate planar technologies in the very near future. The improved sub-threshold slope is expected to lower power by 50%. Lower thresholds will translate directly to lower supply voltages [14]. While these devices seem magical to digital designers they make analog designers cringe in pain. In an ominous sign of the times, continuous time analog circuits may be made of digitized lengths and widths in order to conform to strict Fin-FET process rules. Stacked short channel Fin-FET devices will replace longer channel lengths. Capacitances increase. Non-transistor elements will arrive as an afterthought if at all.

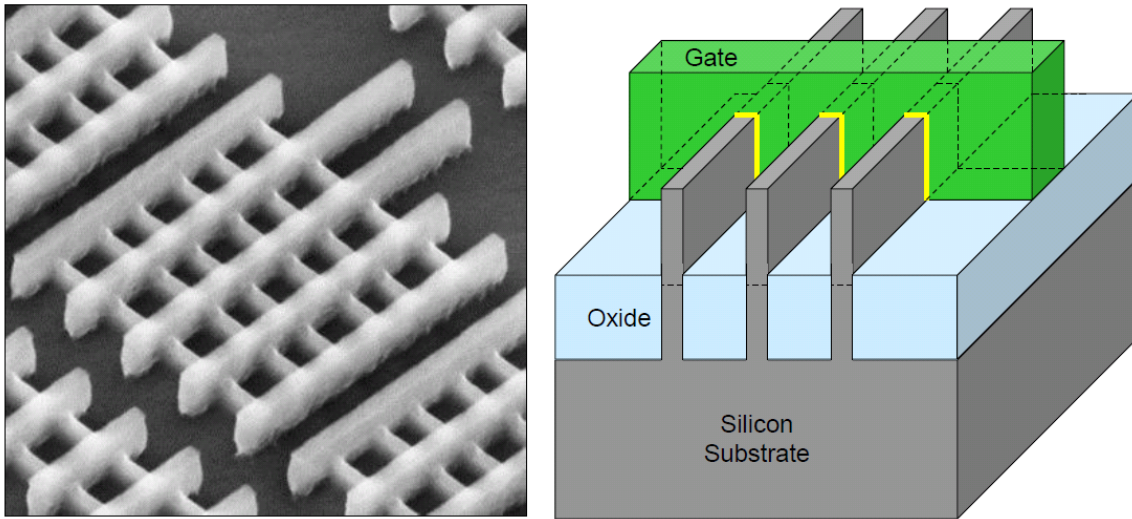


Figure 5: Intel Fin-FET transistors [14]

These issues lead to digitally assisted analog functions. Digitally assisted analog allows for tuning, debug, and test. An integrated power supply solution must follow along with the same trend.

1.7 Fully Integrated Voltage Regulation

Fully integrated voltage regulation implies that a single external source powers a host of internal regulator stages as shown in Figure 4. Research work from the University of California at Berkeley [33] presents this architecture. However, this amount of change limits the adoption of integrated regulation altogether.

Clearly the noise performance of an integrated switching regulation stage prohibits its usage for sensitive analog or mixed signal circuits. Thus, the need for a linear stage after the switching stage remains as shown in Figure 4. Yet other IP could potentially use the main input supply directly if the level aligns with their needs (low speed I/O, SMBUS, I²C, etc.).

Done properly, many advantages exist for the proposed system. The system board eliminates a host of external components. This reduces cost and improves the form factor.

Input power current requirements can reduce if the external supply level increases while keeping the final load current constant. Internal regulator inefficiencies decrease this benefit, however.

Other benefits include fine-grain power regulation for blocks and reduced accuracy requirements of the main input VDD .

1.8 Summary

In summary, many factors contribute to the rise in demand for an integrated regulator solution.

- Parallelization and further integration increase the number of voltage domains desired.

- Power dominates the priority of server designs as massive datacenters provide processing power to an increasing number of handhelds.
- Physical form factors, packaging, and metal resources cannot keep pace with significantly higher levels of integration.

Integrating the voltage regulator follows the same natural progression that many other functions and features have over the history of VLSI and addresses these problems.

2 BACKGROUND

2.1 Linear Regulation

A linear regulator, or low dropout regulator (LDO), could potentially serve as an integrated voltage regulator. Analog circuits incorporate one or more linear regulators for the purpose of power supply rejection and voltage transformation already. The linear regulator provides clean power but does not have good efficiency. This is due to the DC current from the input to the output passing through the regulator. Equation (1) provides an expression for the power efficiency of an ideal linear regulator where I_q is the analog current required for the regulator.

$$\eta_{power} = \frac{P_{load}}{P_{total}} = \frac{V_{load}}{V_{DD}\left(1 + \frac{I_q}{I_{load}}\right)} \quad (1)$$

Figure 6 depicts the efficiency of an ideal linear regulator across load according to (1).

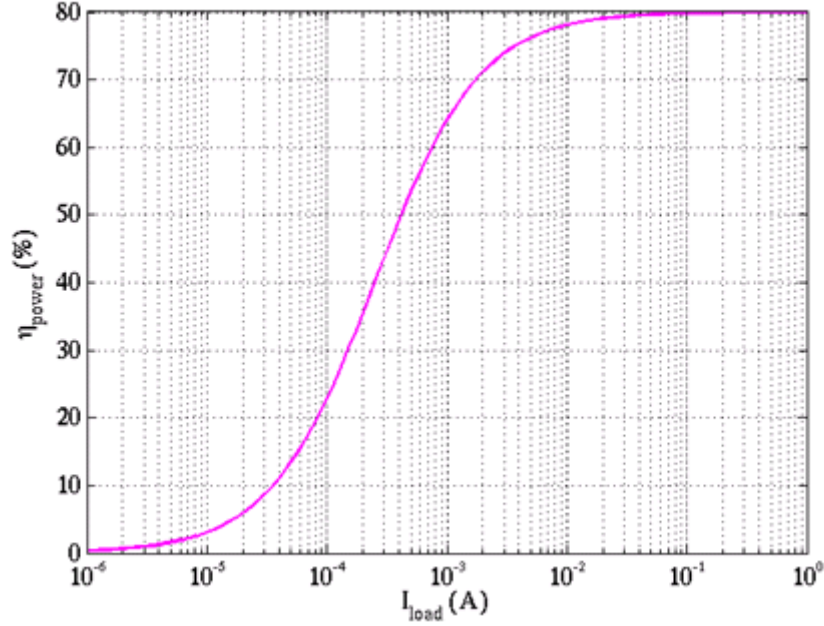


Figure 6: Power efficiency versus output current of a linear regulator with $V_{DD} = 1.5$ V, $V_{load} = 1.2$ V, and $I_q = 100$ μ A

The theoretical maximum efficiency of linear regulator asymptotically approaches V_{out}/V_{in} as the load power takes over. This makes the LDO unsuitable for digital logic load profiles where the DC current reaches high levels.

Also the linear regulator is unsuitable where the input is high voltage/low current and the output is low voltage/high current.

However, there are advantages to the linear regulator. The overhead required to operate a linear regulator can be very small, in the order of a few microamps, which may imply that at low loads a linear regulator can be utilized. Linear regulators can also provide a lower output noise since there is no switching behavior.

2.2 1897: Nikola Tesla's Currents of Ordinary Character

Nikola Tesla wrote a patent for one of the earliest switching power supplies in New York over a century ago:

The invention upon which my present application is based is an improvement in apparatus for the conversion of electrical currents of ordinary character-such, for instance, as are obtainable from the mains of municipal electric light and power systems and either continuous or alternating-into currents of very high frequency and potential.

...

...the said plates being arranged in the manner described, whereby the condensers will be alternately charged in multiple and discharged in series, as set forth. [57]

Tesla's drawing from the patent is shown in Figure 7.

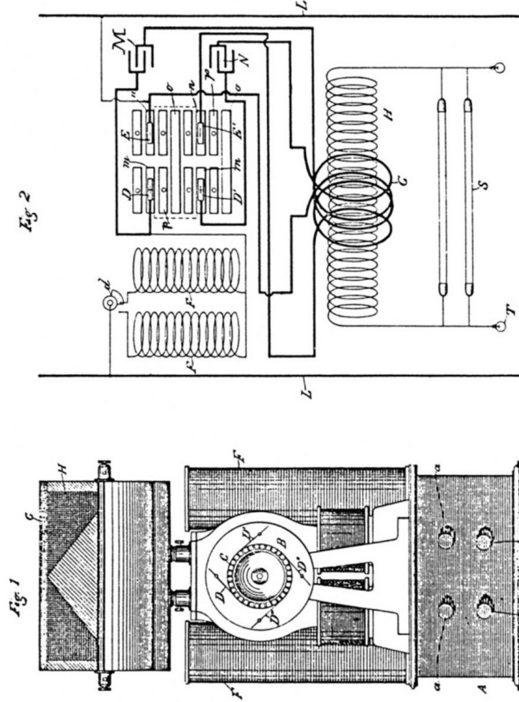
(No Model)

N. TESLA.

APPARATUS FOR PRODUCING CURRENTS OF HIGH FREQUENCY.

No. 583,953.

Patented June 8, 1897



WITNESSES
G. B. Loris.
Edwin B. Apperson.

INVENTOR
Nikola Tesla.
BY
Kerr, Curtis & Page
ATTORNEYS.

Figure 7: Nikola Tesla's "Apparatus for Producing Currents of High Frequency" from 1897 [57]

One hundred fourteen years later Intel [55], Advanced Micro Devices [31], MIT [47], and UC Berkeley [2] explored similar pursuits in the interest of integrating power supplies onto large scale SOC's.

Several recent works have investigated methods to realize a fully integrated voltage regulation (IVR) system.

2.3 Sixteen-Phase Integrated Solution

Belgians Tom Van Breussegeem and Michiel Steyaert published the first integrated switched-capacitor converter achieving high efficiency in 2009 [7]. The design excludes an analog regulation stage and large output capacitor, which results in a compact and efficient design. A multitude of 16 phases combine to form the output, exploiting the fact that integrated regulation allows for a greater number of parallel stages with minimal added cost. The loop is closed by incorporating an amplifier that drives the clock frequency up or down similar to work by Yogesh Ramadass published two years prior [46].

In many ways this work set a reference point for later implementations' performance. The apparently simple system achieved 82% peak efficiency and 0.5% output voltage ripple. Breussegeem and Steyaert formulated a figure of merit for converter voltage ripple. Equations (2) and (3) formulate the ideal ripple of a single capacitive charge pump phase and Breussegeem and Steyaert's figure of merit, respectively.

$$V_{ripple,simple} = \frac{I_{load}}{fC_{out}} \quad (2)$$

$$FOM_{ripple} = \frac{V_{ripple,meas}}{V_{ripple,simple}} \quad (3)$$

They achieved an astounding FOM_{ripple} ratio of 14. Note that in principle the design from Breussegeem and Steyaert should achieve an FOM_{ripple} of 16 since it comprises 16 parallel phases, but in practice the phases may not align perfectly to reduce the ripple.

A comparison of the work by Breussegeem and Steyaert to that of Ma a year earlier [35] clearly illustrates the benefit of integrated regulators over discrete solutions. Integrated voltage regulators can easily divide into many phases where discrete solutions

may be limited at the PCB level. Breusseger and Steyaert's design operated with 5000x less capacitance and achieved 20x less voltage ripple.

While this work exhibits breakthrough performance and full integration, the system control allows for stability without regard to transient performance, sensitivity, or load regulation.

2.4 AMD 32 nm SOI Capacitive Converter

Researchers at UC Berkeley teamed up with Advanced Micro Devices and published a work on yet another fully integrated switched-capacitor converter with valued performance in [31] and again in [33]. The design incorporates a wide 32-phase interleaving strategy to reduce ripple, twice that of Breusseger and Steyaert. The performance, 81% peak efficiency and 0.55 W/mm^2 , compares to that of Breusseger and Steyaert in 2009 except in 32 nm SOI technology. The design lacks a feedback loop to maintain the output by way of frequency or capacitance modulation, which would be necessary in a production design.

Innovative level shifting circuitry and physical planning stand out in the Berkeley work [31]. Integrated systems need a high output power density, W/mm^2 , to take advantage of the integrated regulator benefits. This design achieved a tolerable efficiency at a moderate power density.

Later Le, Sanders, and Alon provide an analytical design methodology for the switching circuits [33]. Perhaps most importantly, the authors mathematically describe the trade-off of efficiency and power density in switched capacitor regulators. Figure 8 depicts this trade-off from the paper.

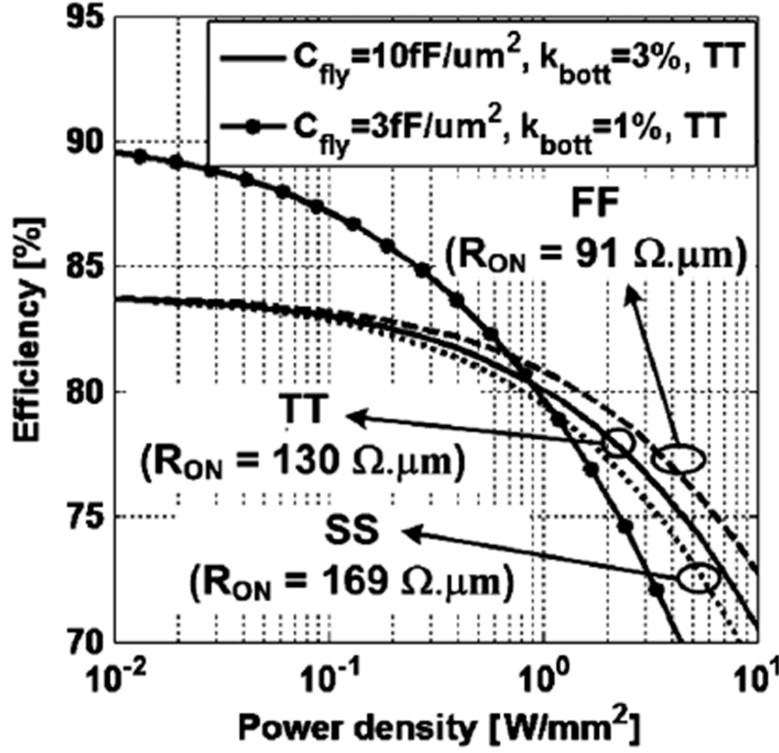


Figure 8: Depiction of the trade-off between efficiency and power density in switched capacitor regulators (from [33])

The graph in Figure 8 illustrates the fact that high efficiency power delivery requires significant area in a switched capacitor solution. Note that the power density limit of the consumption side lies at 2 W/mm². The deliver power density at this same area dips below 75% power efficiency.

2.5 IBM 45 nm Deep Trench Capacitor Converter

IBM engineers published one of the highest quoted efficiencies in integrated regulation systems at the *Symposium on VLSI Circuits* in 2010 [9]. A deep trench capacitor with low parasitic resistance provides a healthy capacitance density of 200 nF/mm², which allows for a maximum delivered power density of 2.2 W/mm². The

design achieves power efficiency of 90% where others peak between 78-83%, thriving on the large capacitance. Figure 9 depicts the trench capacitor and circuit core of the converter.

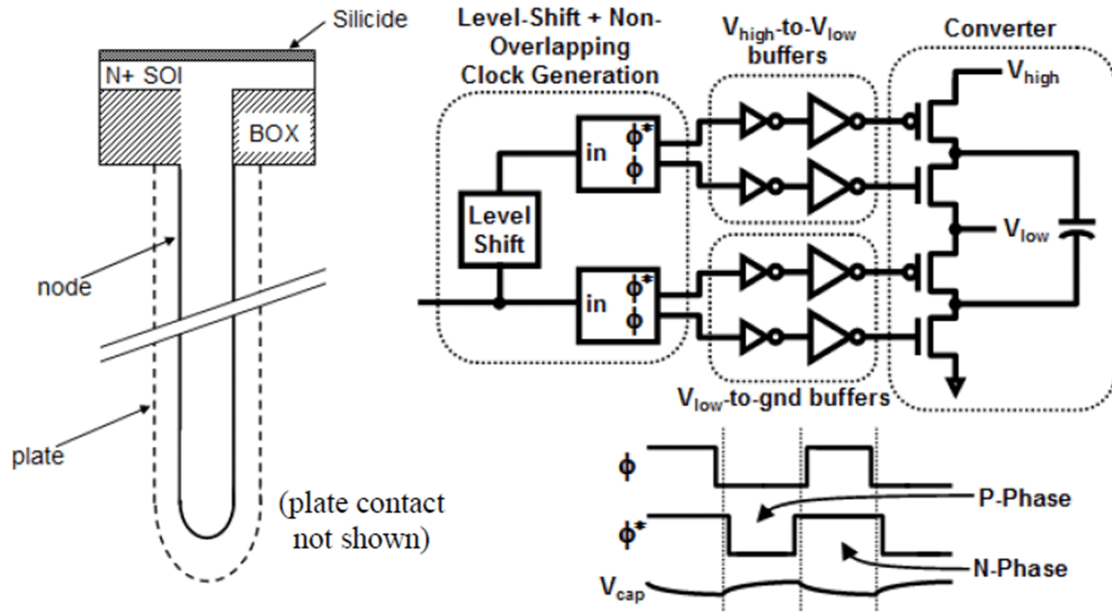


Figure 9: IBM deep trench capacitor converter (from [9])

The IBM converter makes use of dual rail drivers to switch current into and out of the power capacitor. This reduces the parasitic dynamic power of the converter and improves the efficiency.

The design operates in an open loop fashion as published, but the designers mention the possibility of controlling the output regulation by means of clock frequency.

2.6 Conventional Voltage Regulator Down (VRD) Solution

Intel publishes a guide for original equipment manufacturer (OEM) motherboard power supplies for Intel processors called the *Voltage Regulator-Down (VRD) 11.1*:

Processor Power Delivery Design Guidelines [28]. The guidelines serve as a reference point for high performance conventional regulator requirements. Target processors consume up to 180 Watts in their highest power states and have nearly 1400 pins. Even though the module has many high performance cores, one continuous piece of metal supplies up to 185 amps of total current to all of them. Figure 10 illustrates the recommended motherboard power delivery network to the package pins.

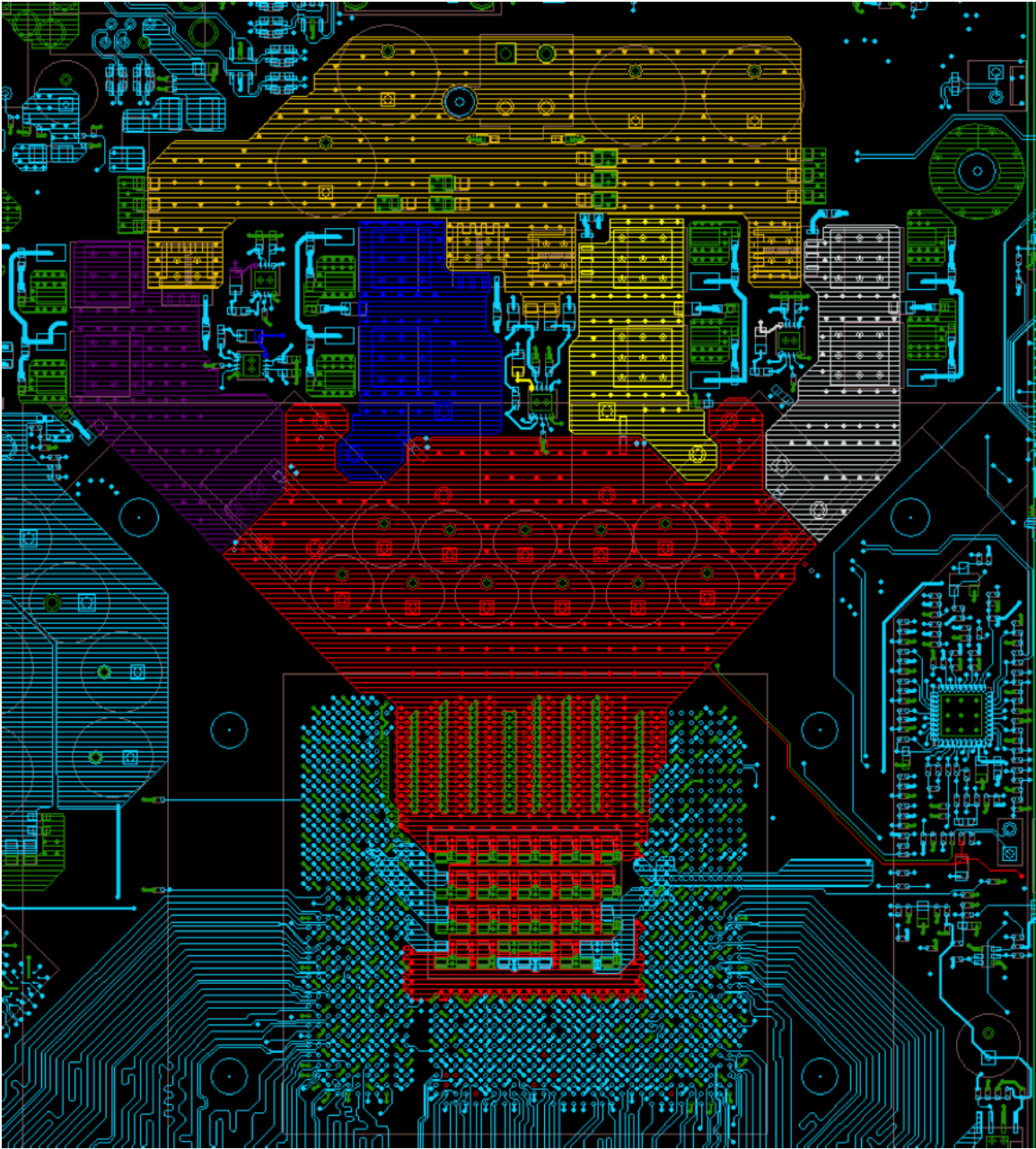


Figure 10: Intel LGA1366 recommended motherboard power routing (from [28])

Large floods and planes of VDD and VSS occupy the area between the processor and the voltage regulator on all 6 layers of the motherboard to reduce parasitic resistance and inductance.

The guideline provides a look into a production design that solves existing power management problems in servers. Many specifications are left up to the OEM, allowing for end product customization. Drawing conclusions from the guidelines about the design, Table 2 enumerates an expanded voltage regulator specification.

Specification	Symbol	Conventional
DC input voltage	$V_{DDB_{DC}}$	12 V
DC output voltage	$V_{REG_{DC}}$	1.1 V
Nominal duty cycle	D	9.2%
Total multi-phase current	I_{total}	180 A
Per phase capacitance	C_{ϕ}	200 μ F
Power efficiency	η_{pwr}	70-80%
Load line resistance	R_{ll}	0.8 m Ω
Output voltage tolerance	$V_{REG_{TOL}}$	220 mV
VID transition slew rate	$S_{\Delta VID}$	10 mV/ μ s
Voltage ripple	V_{ripple}	10 mV
Number of phases	n	4
Max phase current capacity	$I_{avephase}$	45 A
Source area	A_{src}	4418 mm ²
Load area	A_{ld}	450 mm ²
Area ratio	A_{src}/A_{ld}	10x
Switching frequency	f_{sw}	750 kHz
Inductance	L	470 nH
Phase delta-I	Δi	2.2 A
Charge current slope	m_1	23.2 MA/s
Discharge current slope	m_2	2.3 MA/s
Source power density	SPD	0.045 W/mm ²
Source current density	SCD	0.041 A/mm ²
Feedback latency	τ_{fb}	1-2 μ s

Table 2: Conventional regulator design specifications

Note that the output voltage tolerance in Table 2 includes the load line voltage according to R_{ll} ($V_{REG} = VID - I_{REG} * R_{ll}$), a 50 mV allowable excursion above the VID level during transients, and a 19 mV tolerance in the DC load line voltage. Intel also

provides a single bit of feedback from the processor to the voltage regulator, which indicates that the CPU is in a low power state. The VID code is a digital word that represents an output voltage level request from the CPU to the regulator.

2.7 Intel Thin Film Ferrite Multiple Chip Module

Many of the proposed integrated technologies incorporate magnetic energy storage rather than capacitive. Intel researched the deposition of a thin film $\text{Ni}_{80}\text{Fe}_{20}$ layer on top of the uppermost metal in order to construct a high quality power inductor [38] within a semiconductor die. The initial incarnation of this system was a multiple chip module (MCM) solution. The authors tout a tremendous current density of 8 A/mm^2 , 76% peak efficiency, and whopping 400 Amps of output current capacity [17]. The solution, shown in Figure 11, booted and ran a Microsoft Windows stress test for several hours. Clearly the long-term strategy integrates the separate power magnetic die and the processor into the same die.

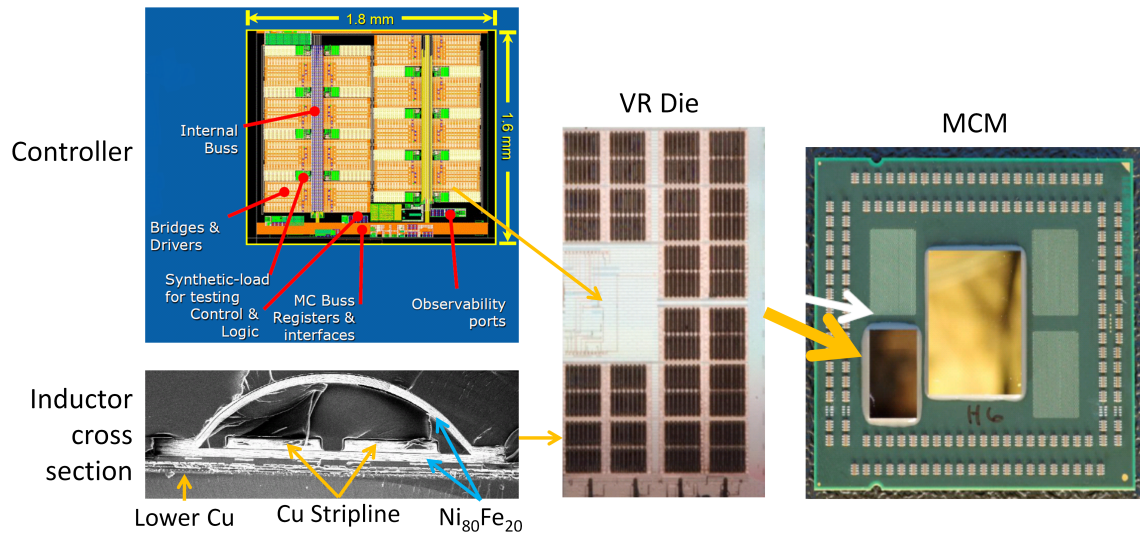


Figure 11: Intel thin film $\text{Ni}_{80}\text{Fe}_{20}$ magnetic solution (from [17])

2.8 Intel Package Inductor MCM

Yet another work put forth by Intel circuit engineers incorporates inductors onto the package. Discrete inductors comprised one solution. The other instance exploited package trace mutual inductance for energy storage as a comparison [53]. Figure 12 depicts the package with mutually coupled traces for inductors.

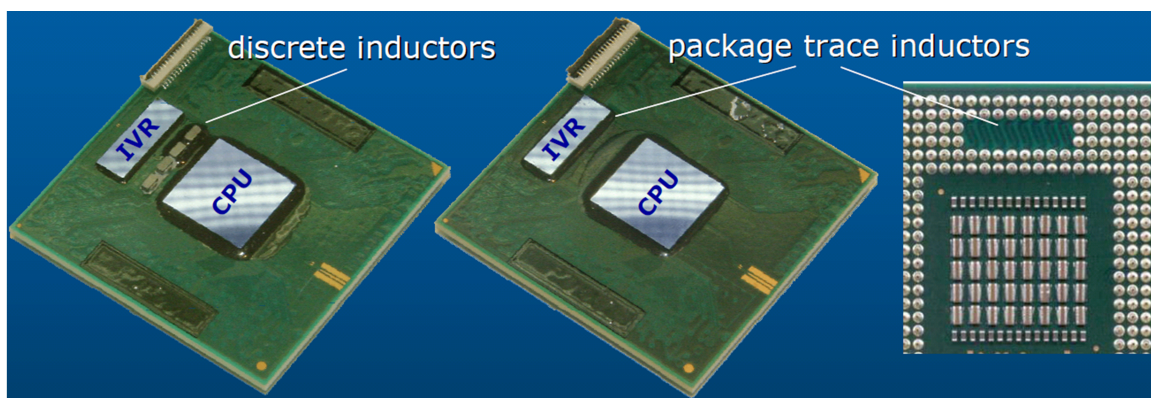


Figure 12: Intel package inductor solutions (from [53])

Package trace inductors do not provide a very good inductor for use in a high power buck converter. These package traces provide a maximum of 1-5 nH of inductance. This pushes the frequency of the buck to a significantly higher frequency, which can lead to lower efficiencies when switching from a higher voltage. In this case the buck converts 3.3 V into roughly 1.0 V. Transistors that support 3.3 V have low performance in terms of power required to deliver current efficiently. As geometries shrink the 3.3 V devices may be eliminated.

2.9 Summary

Technologies from the late 19th century onward have provided solutions to various power conversion problems. Table 3 summarizes the key technologies along with their advantages and disadvantages.

Year	Description	Organization	Advantages	Disadvantages
	Linear regulator		<ul style="list-style-type: none"> • Simple • Low noise output • CMOS construction • 100% integrated 	<ul style="list-style-type: none"> • Efficiency less than V_{out}/V_{in} (ex: 50% for 2V-1V) • Efficiency decreases with power output
2009 [7]	16-phase switched capacitor converter	Breussegem & Steyaert	<ul style="list-style-type: none"> • Voltage ripple FOM of 14 • CMOS friendly 	<ul style="list-style-type: none"> • 2 mW/mm² density • Narrow voltage range support
2010 [31]	AMD 32 nm series-parallel SC converter	AMD/UC Berkeley	<ul style="list-style-type: none"> • Improved range support • 80% peak efficiency 	<ul style="list-style-type: none"> • Low efficiency across load range • 0.9 W/mm² power density
2010 [9]	IBM 45 nm flying trench capacitor	IBM	<ul style="list-style-type: none"> • 90% peak efficiency • 2.2 W/mm² power density 	<ul style="list-style-type: none"> • Requires deep trench capacitor process • Only supports 2:1 input to output ratio
2009 [28]	Motherboard buck regulator	Intel	<ul style="list-style-type: none"> • Existing solution with long history of production success • Shorter runway to production 	<ul style="list-style-type: none"> • 70-80% efficiency • $10 \times A_{src}/A_{ld}$ • 1-2 μs feedback latency • Unsuitable for creating many rails • Physically large
2010 [17]	Thin film ferrite MCM	Intel	<ul style="list-style-type: none"> • Supports 8 A/mm² • Could be CMOS single process, stack, or MCM 	<ul style="list-style-type: none"> • Significant cost and effort to implement inductor • 76% peak efficiency
2010 [53]	Package trace inductor buck converter	Intel	<ul style="list-style-type: none"> • Lacks complex process steps • 85% efficiency 	<ul style="list-style-type: none"> • High inductor DC resistance • High input voltage supply

Table 3: Summary of prior power conversion solutions

3 SWITCHED-CAPACITOR SOLUTIONS

The following chapter describes research in the design and implementation of integrated switched-capacitor voltage regulators.

3.1 Generalized IVR Control System

Most of the recent research of integrated voltage regulators focuses on the energy storage components (capacitors and inductors) and topologies thereof. This chapter focuses on simulation methods for complex switching systems, such as a capacitive converter, for use in a highly integrated system.

Figure 13 depicts the general IVR system for both capacitive and inductive regulators. A capacitor and resistor in parallel model the load. A series resistor represents the metal distribution grid between the regulator phases and the load itself. Capacitive or inductive phases operate as the actual conversion mechanism.

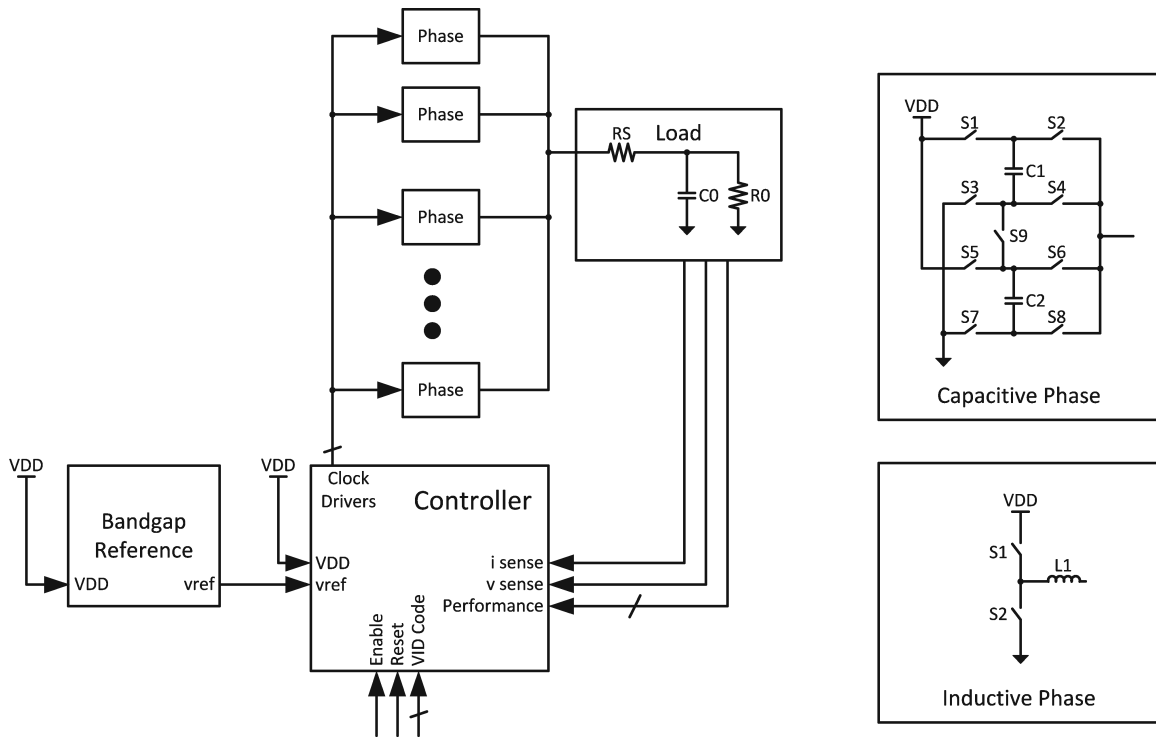


Figure 13: Switched-capacitor integrated regulation system.

The input power supply and reference circuit form feedforward paths to the output. The Voltage IDentification (VID) input bus sets the DC output level. The performance message bus provides a communication channel between the load circuitry performance counters and the regulator controller itself with low latency.

Feedback paths include both output voltage and current ports. Sensing the output voltage at the load eliminates DC error due to R_S . Circuitry monitoring the energy storage element (capacitive or inductive) estimates the output current, ideally without significant loss, in the primary output current path.

3.2 Converter Modeling

A mathematical model of the converter phase itself aids in the design of the controller. An analytical hand model extends to an autonomous MATLAB model for loop dynamics research.

3.2.1 Manual Solver

One particular switched-capacitor converter phase includes 2 capacitors and 9 switches as shown in Figure 14.

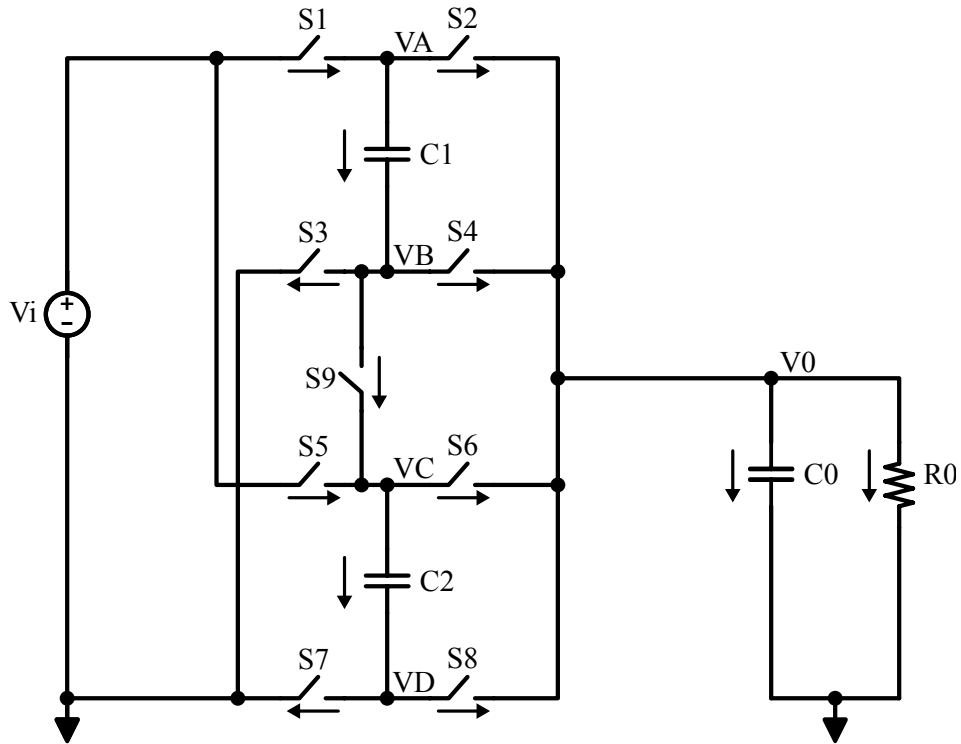


Figure 14: Single-cell circuit schematic [33]

In 2:1 conversion mode the converter operates under two clock phases. The first phase closes switches S_1 , S_4 , S_5 , and S_8 , charging C_1 and C_2 in series. The second

phase closes S2, S3, S6, and S7, discharging C1 and C2 in parallel into C0. The cell also supports 3:1 and 3:2 conversion ratios by changing the switching order.

Kirchhoff's current law (KCL) equations (4) through (8) provide a mathematical model for Figure 14 that combines with higher level clocking to form a regulator. R_n denotes the resistance of switch S_n and varies with time according to the input clocks.

$$V'_{C1} = \frac{1}{R_1 C_1} V_i + \left(\frac{1}{R_1 C_1} + \frac{-1}{R_2 C_1} \right) V_A + 0 \cdot V_B + 0 \cdot V_C + 0 \cdot V_D + \frac{1}{R_2 C_1} V_0 \quad (4)$$

$$V'_{C1} = 0 \cdot V_i + 0 \cdot V_A + \left(\frac{1}{R_3 C_1} + \frac{1}{R_4 C_1} + \frac{1}{R_9 C_1} \right) V_B + \frac{-1}{R_9 C_1} V_C + 0 \cdot V_D + \frac{-1}{R_4 C_1} V_0 \quad (5)$$

$$V'_{C2} = \frac{1}{R_5 C_2} V_i + 0 \cdot V_A + \frac{1}{R_9 C_2} V_B + \left(\frac{-1}{R_5 C_2} + \frac{-1}{R_6 C_2} + \frac{-1}{R_9 C_2} \right) V_C + 0 \cdot V_D + \frac{1}{R_6 C_2} V_0 \quad (6)$$

$$V'_{C2} = 0 \cdot V_i + 0 \cdot V_A + 0 \cdot V_B + 0 \cdot V_C + \left(\frac{1}{R_7 C_2} + \frac{1}{R_8 C_2} \right) V_D + \frac{-1}{R_8 C_2} V_0 \quad (7)$$

$$V'_{C0} = \frac{1}{R_2 C_0} V_A + \frac{1}{R_4 C_0} V_B + \frac{1}{R_6 C_0} V_C + \frac{1}{R_8 C_0} V_D + \frac{1}{C_0} \left(\frac{1}{R_0} + \frac{1}{R_2} + \frac{1}{R_4} + \frac{1}{R_6} + \frac{1}{R_8} \right) V_0 \quad (8)$$

The integration of equations (4), (6), and (8) in time determines the state change at each time step. V'_{C0} , V'_{C1} , and V'_{C2} can, in turn, determine the internal node voltages with the help of equations (5) and (7). These equations are solved in MATLAB matrix form for faster performance. This structure also allows expansion of the model to include many single cell converters in parallel.

Figure 15 depicts the step response correlation of the converter mathematical model to a Cadence Spectre circuit simulation. The mathematical model and circuit simulations correlate within 50 μ V.

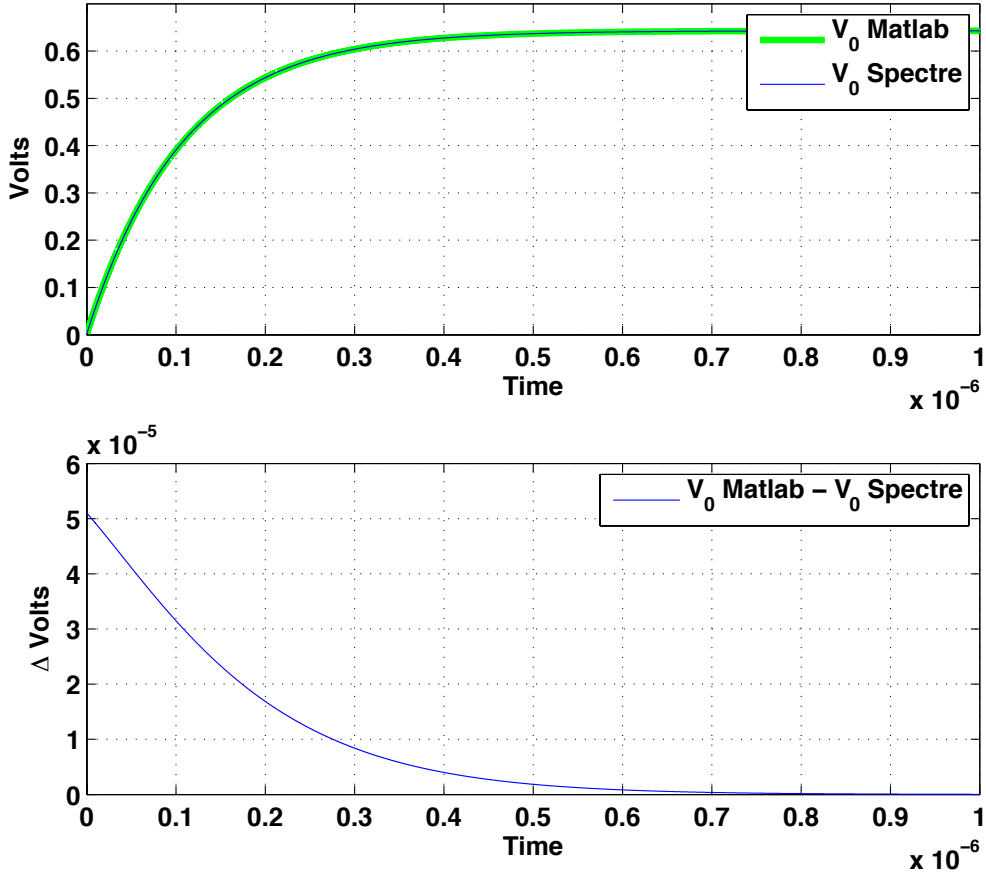


Figure 15: Comparison of single-cell DC simulations in MATLAB and Cadence Spectre

Once the preceding mathematical solution correlates, the resistors can turn into clocked voltage controlled resistors to mimic the behavior of switches. Figure 16 depicts the internal capacitor voltage, V_{C1} , and the output voltage, V_0 . Fixed frequency clocks drive converter switches at 1 GHz in 2:1 mode where the DC output level targets one-half the input supply. Under this load and configuration, the internal cell capacitor discharges into the load every other clock phase and recharges to the input voltage in the opposing phase. The converter output voltage averages to approximately 0.9 V with an input of 1.8 V as targeted.

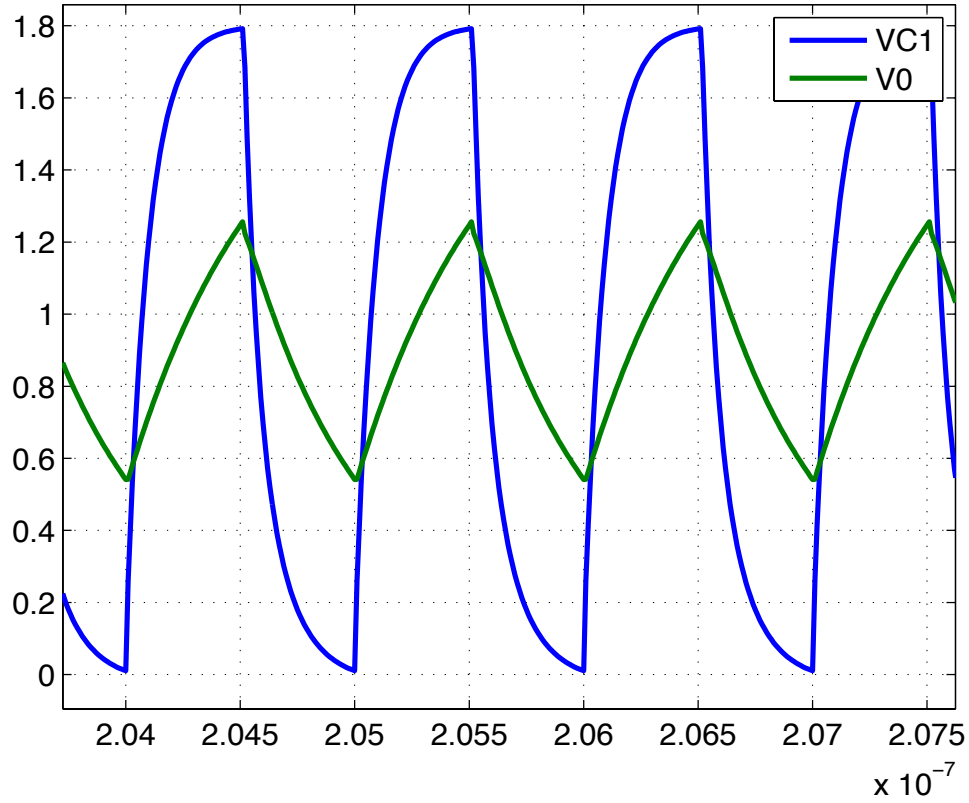


Figure 16: Single-cell converter mathematical model clocked in 2:1 conversion mode

3.2.2 MATLAB and Circuit Cosimulator

Solving KCL equations and matrices for each converter design provides insight into the internal behavior of the converter phase. However, this method requires too much time. An automatic MATLAB hybrid simulator incorporates a user written netlist similar in format to a SPICE netlist. A custom network solver parses this netlist and mechanically produces a sparse matrix of conductances, node voltages, and branch currents automatically at run time.

Resistors, capacitors, and voltage sources translate into matrix “stamps” as in SPICE [39]. These stamps combine systematically in conductance (\mathbf{G}), voltage (\mathbf{V}), and current (\mathbf{I}) matrices. The simulator then integrates $\mathbf{GV} = \mathbf{I}$ in a transient simulation.

Figure 17 illustrates the interactions of the enhanced modeling method.

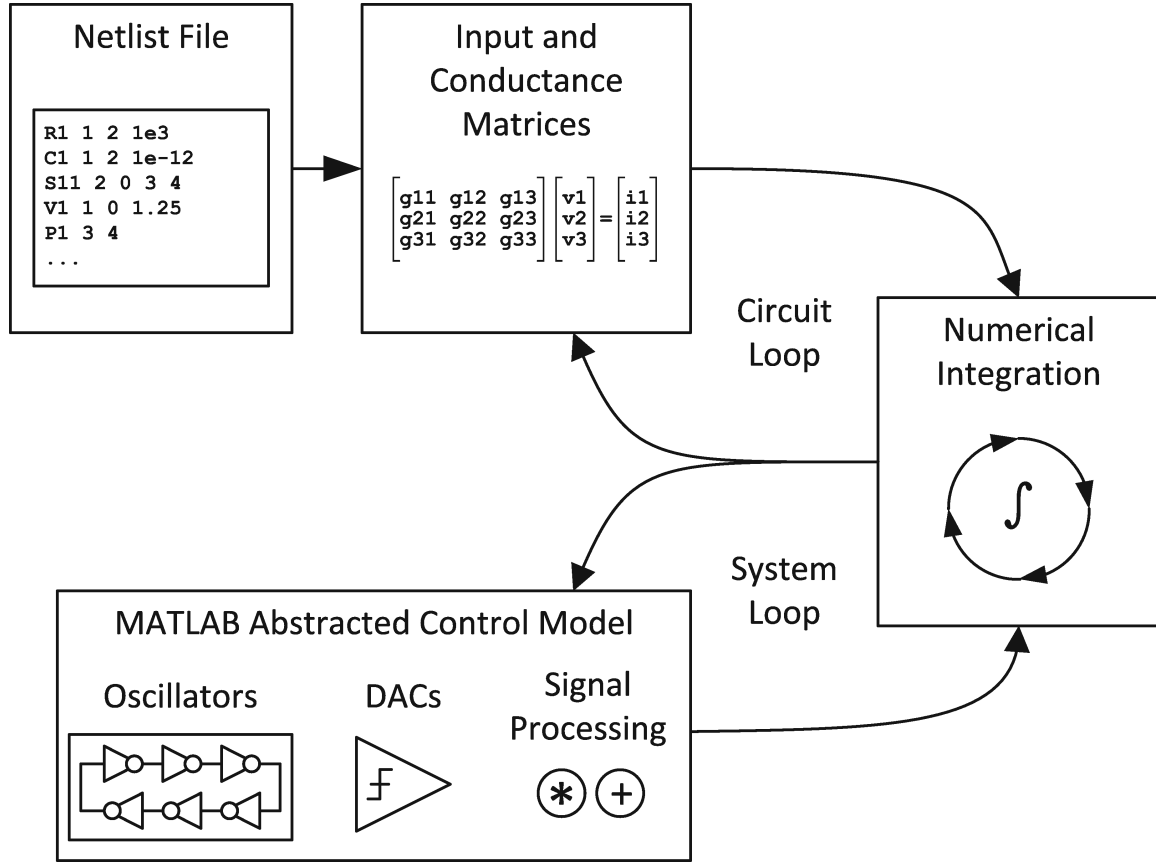


Figure 17: Circuit and behavioral element interaction in the automatic MATLAB hybrid model

Solving key circuit topologies using the netlist method allows for accurate modeling in the converter phase while the rest of the architecture simulates in a behavioral sense. The netlist parser understands special elements that tie matrix elements to MATLAB variables outside of the high accuracy circuit loop. Thus, a high level

MATLAB module can provide input to the circuit solver. This method proves beneficial for solving switched-capacitor power converter architecture problems.

The first crude feedback loop implements a behavioral VCO, which responds proportionally in frequency to the output of an analog amplifier. Figure 18 illustrates this concept, similar to work published by Breussegem in 2009 [7].

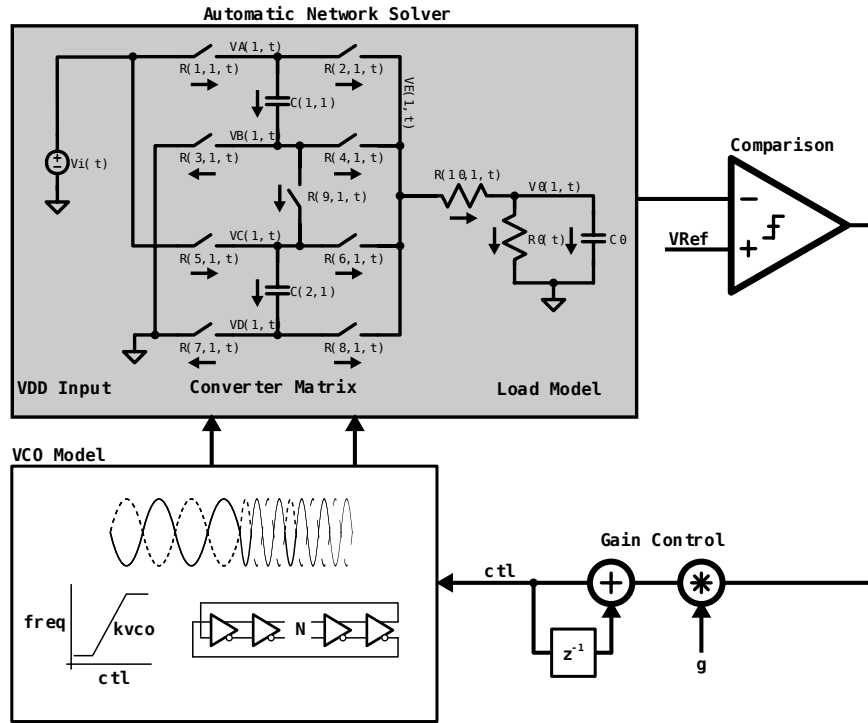


Figure 18: A simple closed loop capacitive converter model

The gain of the VCO and the frequency of comparisons affect the stability and overall output impedance of the supply. Figure 19 shows the regulated output under constant load.

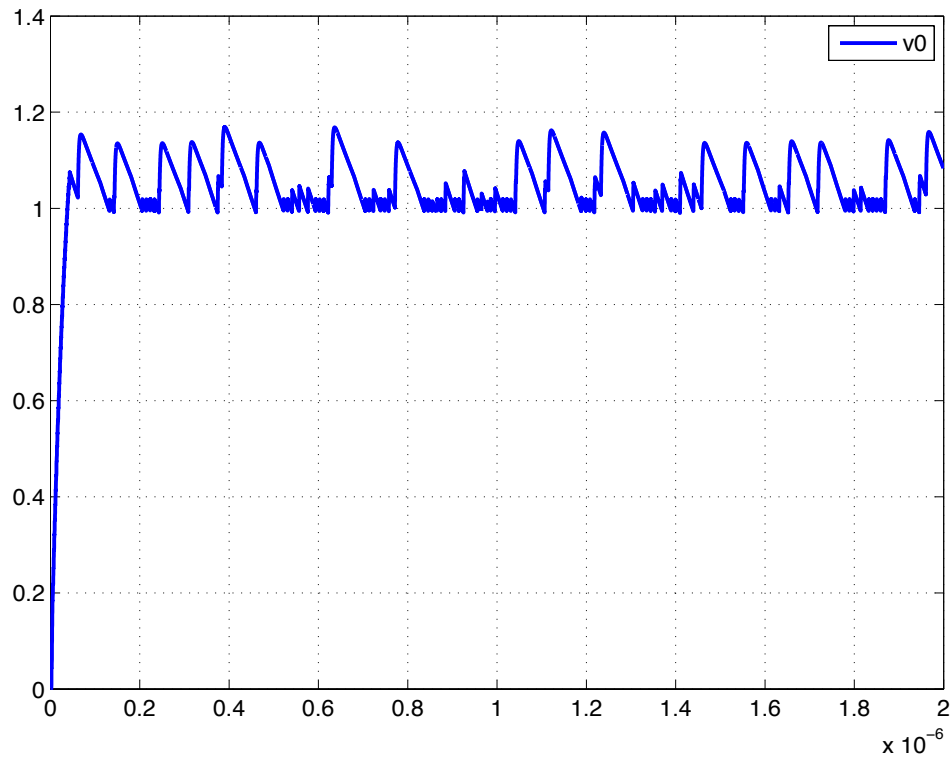


Figure 19: Output voltage versus time of the MATLAB model in Figure 18

The control methodology, albeit crude, needs optimization. Clock frequency alone modulates the output level. However, the simulator finds a solution quickly and accurately.

The cosimulator code reads the listing in Figure 20 to implement the converter phase implemented in Figure 18 and Figure 19.

```

v1 vi 0 1.8
zck1 ck11 0 ck11
zck2 ck12 0 ck12
s11 vi va1 ck11 0 1
s21 va1 ve1 ck12 0 1
c11 va1 0 25e-12
re1 ve1 v0 1
* Output Load
r0 v0 0 3600
c0 v0 0 100e-12

```

Figure 20: Input listing to MATLAB cosimulator for switched-capacitor development.

The code describes a phase in terms of the switches, capacitors, and hooks into the behavioral world. The “z” element, in lines 2 and 3, represent a link to variables outside of the network matrix, which are controlled behaviorally in MATLAB. In this case the VCO model, a behavioral one, drives clocks ck11 and ck12. The “s” element defines a voltage controlled resistance, or switch, with piece-wise linear resistance.

This framework allows infinite flexibility in the converter phase design while maintaining abstraction in more complex modules that comprise the feedback mechanism in the loop.

3.3 Switched Capacitor Usage

Switched capacitor converters do not work well with large DC currents because the switching power becomes prohibitively large or they need a large capacitance. This

limits the usage of switched-capacitor converters for large-scale digital circuits. With sufficiently large capacitance densities and small voltage islands, switched capacitor converters may prove to be a viable solution.

However, the biggest drawback in using a switched capacitor solution is the inability to generate a continuous range in the output voltage. A switched capacitor solution peaks in efficiency at very specific ratios of the input power supply, given that charge is being redistributed fractionally to the capacitors. Operation outside of those fixed ratios incurs a significant loss. In order to circumvent this limitation, some switched capacitor architectures implement complex restructuring algorithms to stitch the discontinuous operating points together. Other solutions incorporate parts of a buck converter to cover these areas.

Analog loads work well with switched supplies as long as architecture is amenable to a single voltage point, unlike digital loads with many different power states.

Nonetheless, the preceding simulation environment proves useful for future complex control models, where high level MATLAB code co-simulates with simple circuit networks.

4 BUCK ELEMENTS

The switched-capacitor cannot provide a continuum of output voltages without high loss. Inductive switching regulators provide a more efficient solution. The simplest form of inductive switching regulator, and the most amenable to integrating, is the basic buck.

The buck voltage regulator employs all of the basic circuit elements. The load behaves predominantly as a resistor. A series inductor integrates the input voltage waveform into a current. A capacitor filters the voltage output to provide a clean source to the load. Inductor and capacitor performance play a crucial part in the story of integrated voltage regulators. This chapter walks through the pertinent components available today for creating a buck regulator that powers a high performance microprocessor. The first items on the table for discussion include the inductor, capacitor, and power FETs.

4.1 Inductor

The most important and problematic element in the system is the inductor. The inductor capabilities, form factor, AC/DC loss, and inductance itself set the design into motion. Inductors come in many shapes and sizes, from parasitic bond wire inductors with 3-5 nH to discrete ferrite core inductors with 100-500 μ H inductance.

Most processor power supplies incorporate several high capacity discrete inductors as shown in Figure 21.



Figure 21: Vishay inductors for high current motherboard power supplies

These high power inductors support currents as high as 15-36 A. Inductances range from 470 nH to 10 μ H. The DC resistances can be as low as 1.6 m Ω .

4.1.1 Integrated Magnetic Power Inductors

On-die, or integrated, inductors are common for LC oscillators in high performance phase-locked loops (PLLs). However, these inductors do not perform well in power applications due to their low performance for high currents. PLLs operate reasonably well given this limitation. They operate at very high frequencies and power efficiency comes second to phase noise performance. That being said, in the last decade a lot of innovation took place in the realm of thin-film integrated power inductors with magnetic material in the windings. Two primary structures for on die power magnetics exist: mutually coupled strips and magnetic yolks. Both have advantages and disadvantages. One is copper surrounded by magnetics. The other is magnetics surrounded by copper.

In the mutually coupled strip, shown in Figure 22, copper wiring carries current parallel to the die and the magnetic field spirals around the wires. Multiple laminations of magnetic material encase the copper wiring to achieve higher inductance. All of the complex processing required to form this structure lies in the magnetic laminations and magnetic vias that encompass the simple copper wires. This structure possesses the

advantage of low resistivity of the copper wire, given the lack of current-carrying vias. However, fabrication of the complex magnetic structure can be costly.

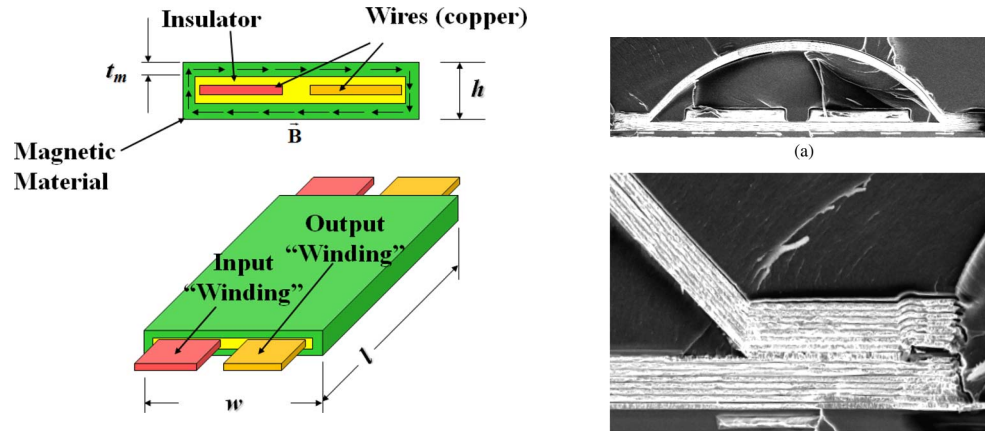


Figure 22: Thin-film strip-style power inductor (from [38])

The magnetic yolk, or coupled solenoid, sends current through a spiral and allows the magnetic field to circulate parallel to the die plane. The structure, proposed by researchers at Columbia University [57], is shown in Figure 23.

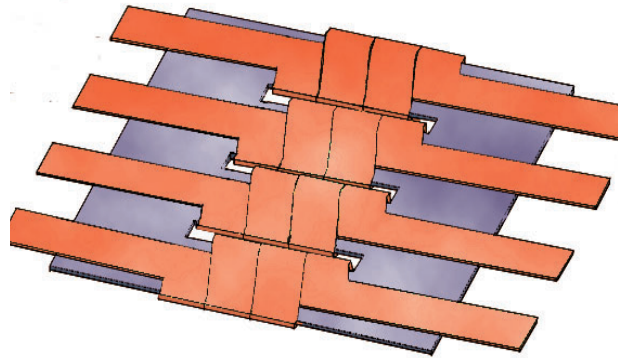


Figure 23: Thin-film yolk-style power inductor (from [57])

The thin-film yolk-style power inductor favors standard CMOS processing steps more than the coupled strips because the complex spiraling occurs in copper rather than magnetic material. This leaves a flat layer of magnetic laminations without magnetic vias, which is a much simpler approach.

All of the thin-film integrated inductors require additional processing steps to incorporate high permeability materials, raising the inductance into a workable range for integrated power conversion. Otherwise air-core structures require too many turns leading to intolerable losses. While these inductors may someday be attainable for a fabless semiconductor company, a lot of development must take place before the transformation can occur. The technology may take several years to mature. In the meantime, opportunity exists to take an intermediate step towards integrated voltage regulators using a discrete inductor available today.

4.1.2 Discrete Inductors

A surge of mobile devices and miniaturization has driven the development of inductors into extremely small form factors that are ideal for integrating onto the surface of the package.

In general, there are three main types of discrete inductors available today which can be mounted on the surface of a package. These include wire wound ferrite core inductors, thin-film ferrite inductors, and air core inductors.

Figure 24 depicts a wire wound ferrite core inductor. This style of discrete component provides a substantial inductance for a very small form factor. The copper wiring provides good conductivity resulting in reasonable DCR levels. Form factors come in standard surface mount sizes down to 1.0 mm x 0.5 mm.

Manufacturers specify the size of discrete surface-mount technology (SMT) components in many different units. The most common unit for SMT components in the United States is the JEDEC inch specification. The JEDEC code specifies the size of the component in one-hundredths of an inch. The first two digits correspond to the length, and the third and fourth digits correspond to the width. A fifth digit may be added for widths much less than one-hundredth of an inch. For example, an 0402 component in the JEDEC standard measures 1.0 mm x 0.5 mm, or 0.039" x 0.020". Similarly, an 01005 component measures 0.4 mm x 0.2 mm, or 0.016" x 0.0079".

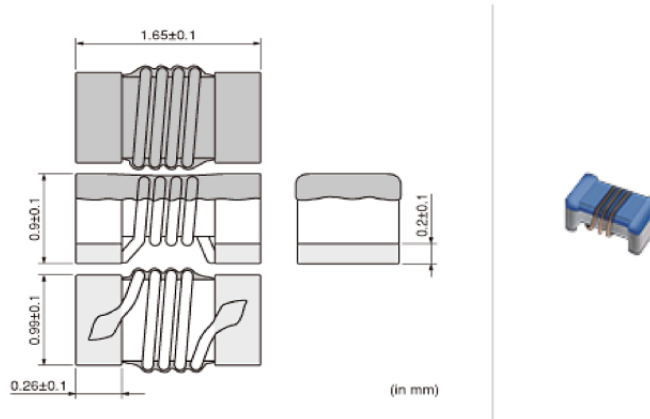


Figure 24: Wire wound ferrite core inductor (from [39])

A selection of 0603 and 0402 wire wound ferrite core inductors from several manufacturers is shown in Figure 25.

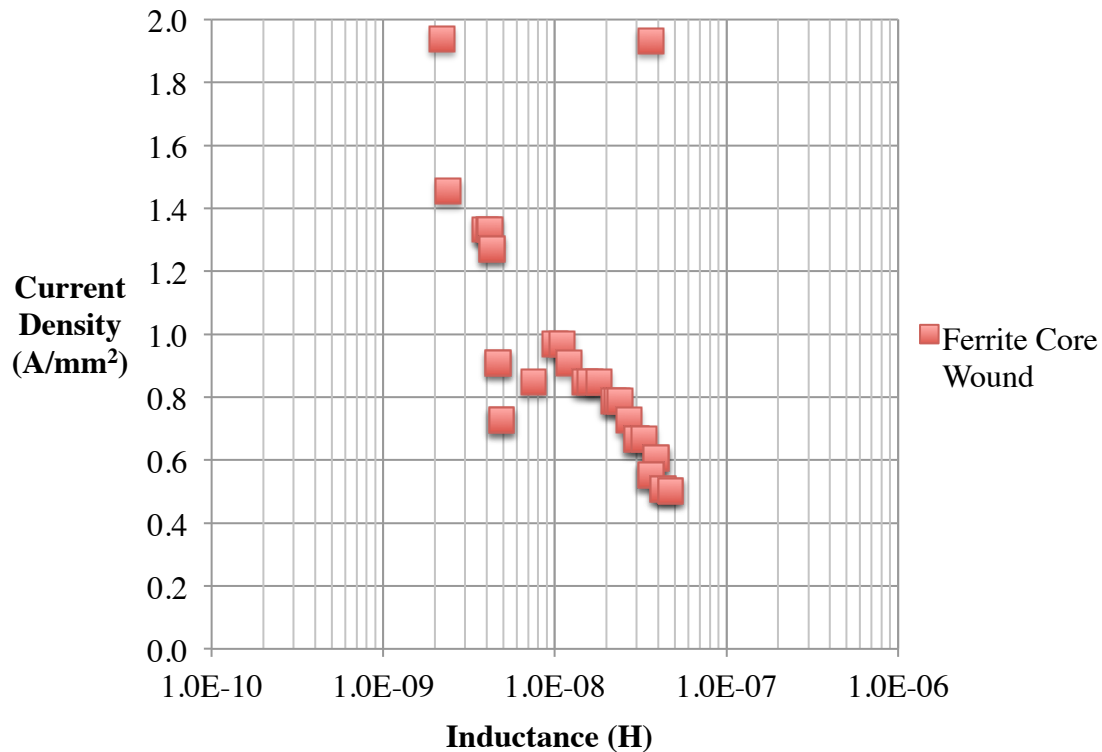


Figure 25: Current density of 0603 and 0402 ferrite core wire wound inductors

The y-axis depicts the current density the inductor supports, and the x-axis represents the inductance at DC. From this the possibility of IVR becomes apparent. Note that the increasing current density requires a significantly lower inductance, which results in a higher switching frequency. At these low inductance levels the dynamic power required for off-die power FETs becomes prohibitively large.

Another interesting device for power conversion is the air core wire wound inductor. Figure 26 shows some of Coilcraft's newest miniature air core inductors [12]. These tiny inductors have lower inductance since there is no magnetic core, but this can be advantageous in the IVR application. The large diameter copper wiring allows for the lowest DC resistance. The air core doesn't suffer from magnetic field saturation near the

operating region. The core loss is also significantly reduced. Reduced DC resistance and lack of magnetic field saturation allows for a significantly higher current capacity.

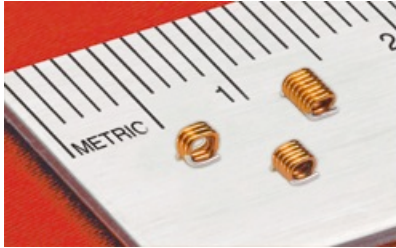


Figure 26: Wire-wound air core inductors (from [12])

The most desirable trait of the air core inductors comes in the form of low DC loss. The DC resistance easily pushes below $10\text{ m}\Omega$ for inductances as large as 27 nH .

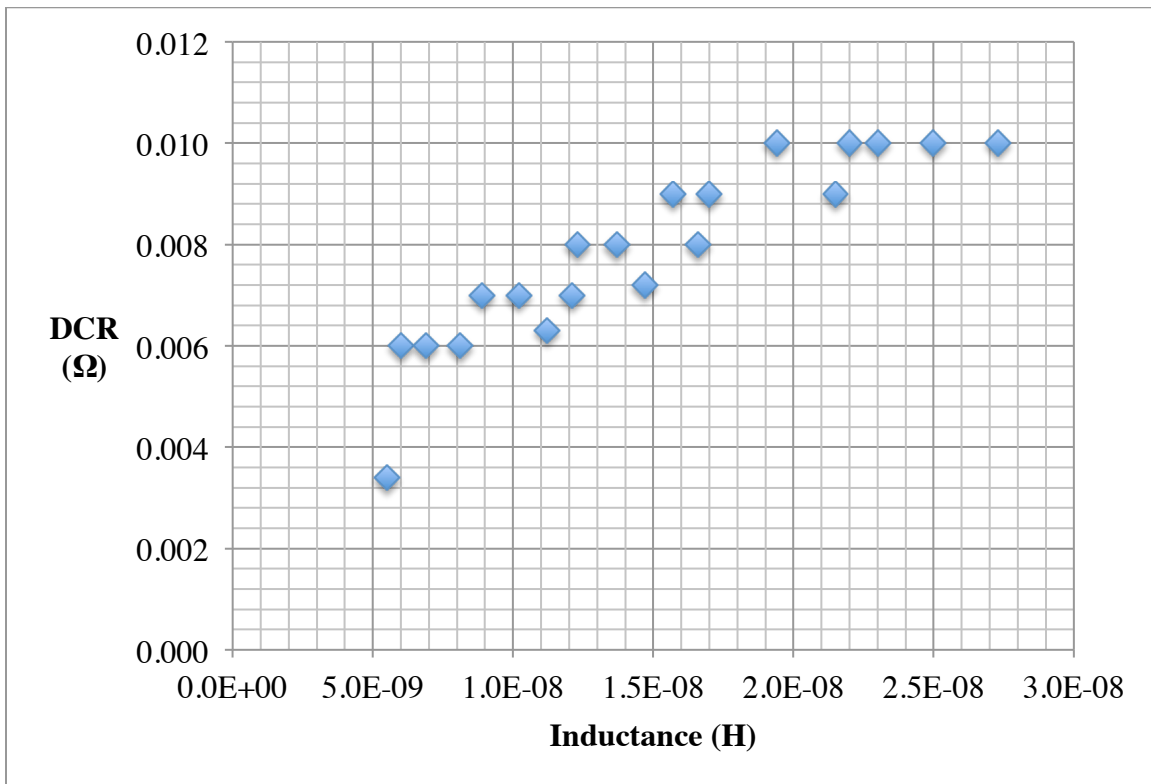


Figure 27: DCR versus inductance for a series of air core inductors

These low DCR levels come with a cost, however. Reaching these levels requires more area.

The third type of discrete inductor is a thin film ferrite structure. This technology is similar to surface mount resistors, except that a magnetic material is used to pattern a coil. The coil produces a magnetic field perpendicular to the mounting surface. These very small inductors yield a very high inductance given the tiny footprint. However, the conductivity in the ferrite material is poor compared to copper, which leads to high losses. These inductors may find a suitable home in very low current targets where the high inductance can be exploited.

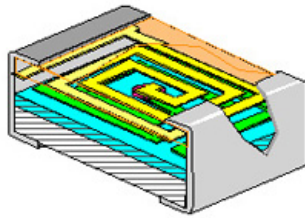


Figure 28: Thin-film ferrite discrete inductor (from [50])

Figure 29 plots current density versus inductance of all three categories of inductors.

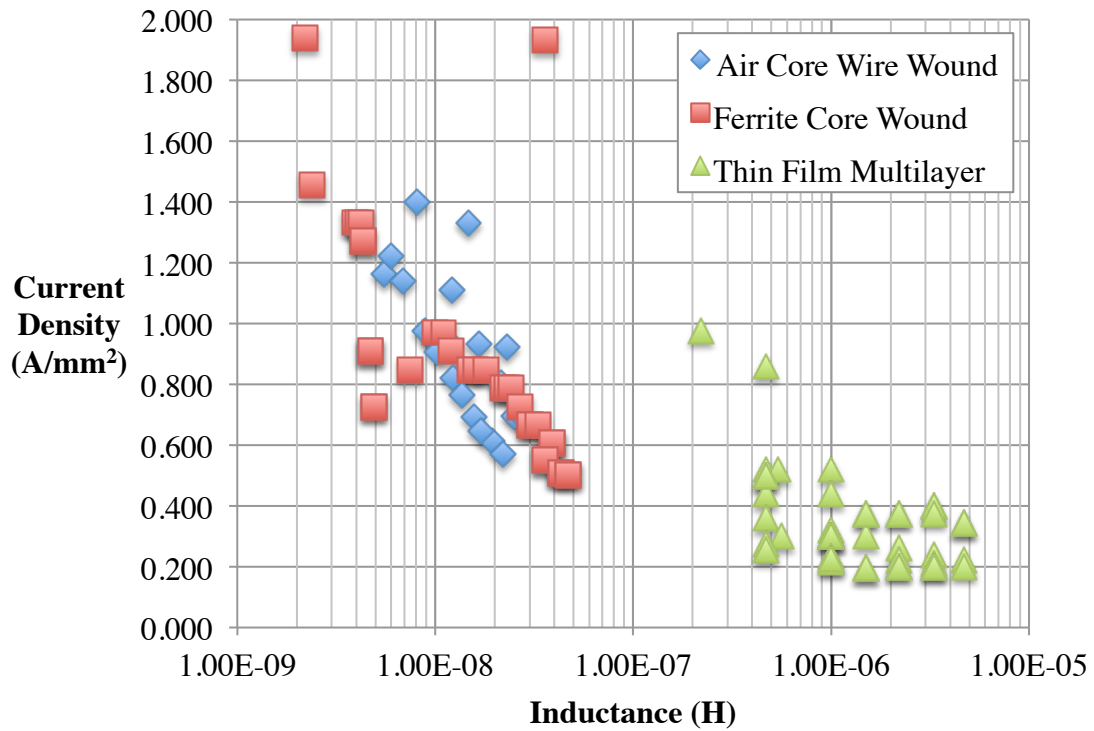


Figure 29: Current density versus inductance for three discrete inductor structures

Note that for a given inductance, air core and ferrite core inductors support comparable current densities. The thin film multilayer types do not support as high of current densities. This becomes even more evident when the current density is plotted versus DC resistance as in Figure 30.

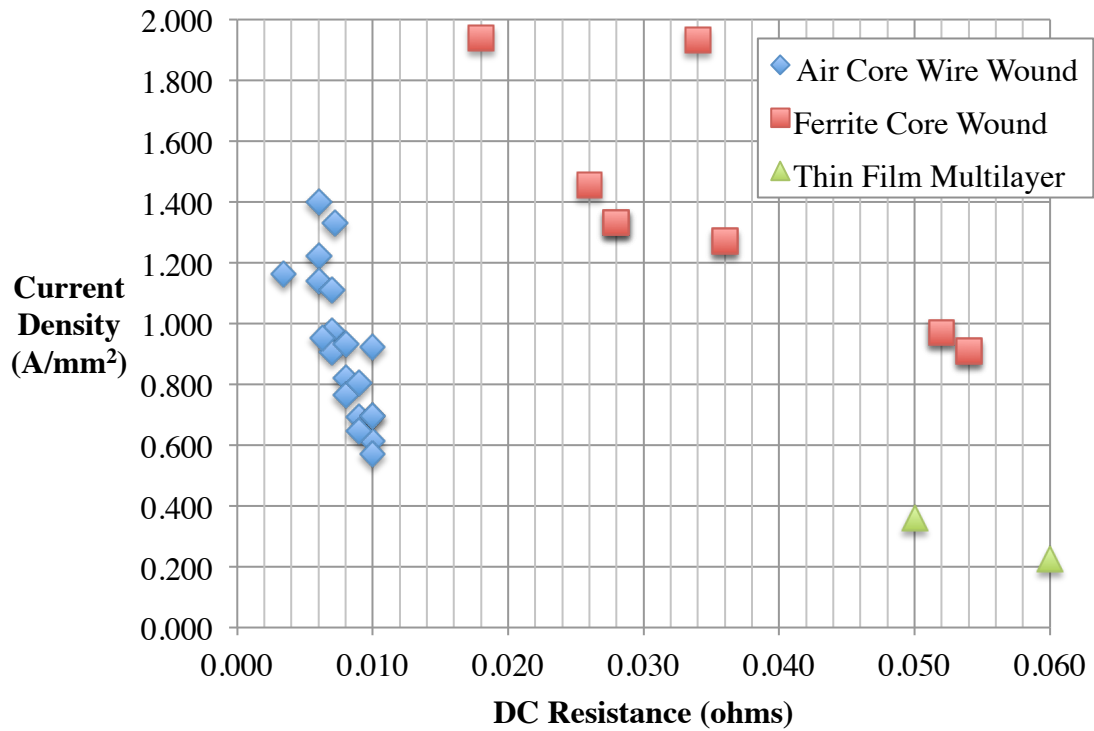


Figure 30: Current density versus DC resistance of inductor types

Note that the ferrite core wound inductors achieve the highest current densities, nearly 2 A/mm².

Another view that helps select an inductor is the conductance versus inductance plot as shown below. This clearly shows the three separate types having very different characteristics.

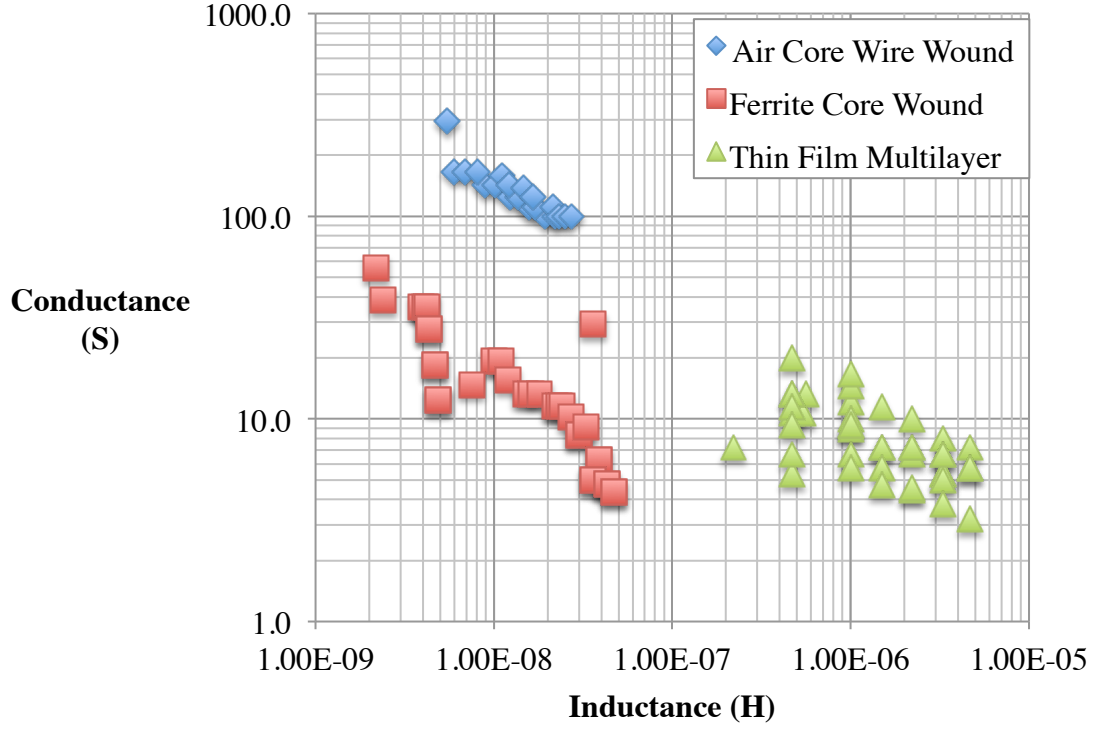


Figure 31: Conductance versus inductance of inductor types

An evolution from the high current toroids to the miniaturized inductors becomes apparent in plotting the conductance versus f_{sw} as in Figure 32. f_{sw} is approximated by assuming a fixed output capacitance, $C_{out} = 10\mu\text{F}$, and oversampling ratio, $x_{os} = f_{sw}/f_0 = 100$.

$$x_{os} = \frac{f_{sw}}{f_0} \quad (9)$$

$$f_{sw} = \frac{x_{os}}{2\pi\sqrt{LC}} \quad (10)$$

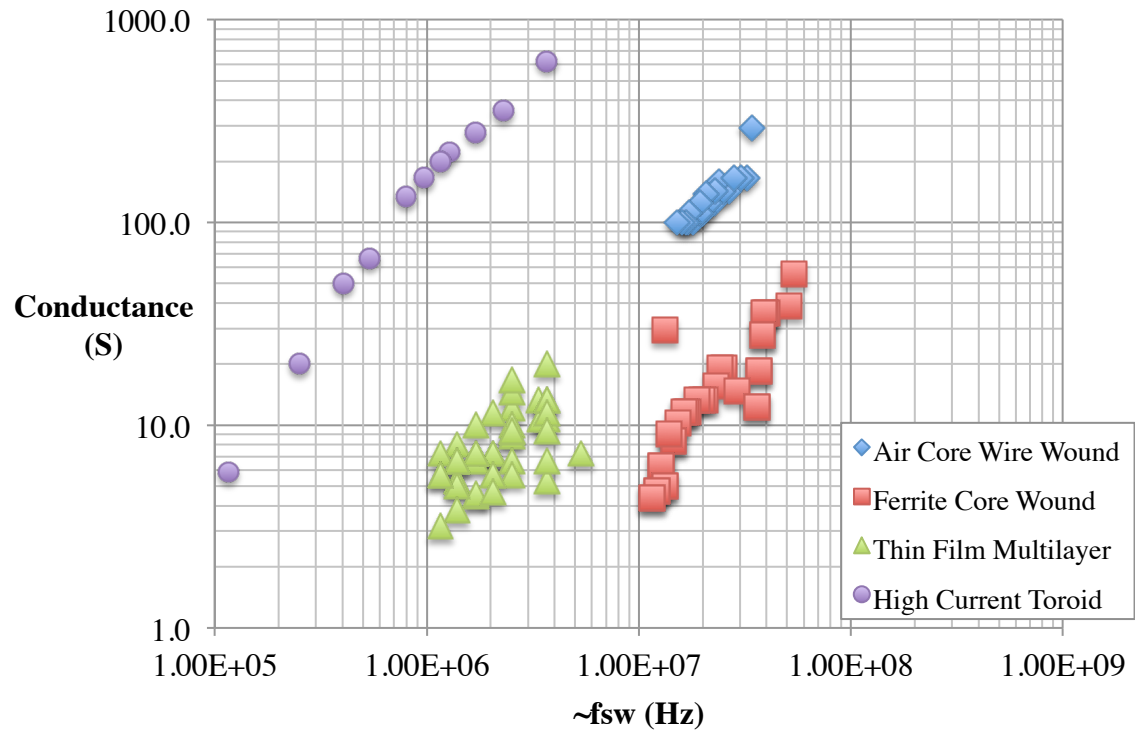


Figure 32: Conductance versus f_{sw} of inductor types

Note that as the inductor shrinks, the switching frequency increases to maintain a comparable system. High current toroid systems operate in the 500 kHz to 1 MHz range on motherboards where the package mountable air core wire pushes as high as 40 MHz.

All of this information will be used in the following chapter to architect a system around one of these inductors.

4.2 Capacitor

Capacitance is needed to filter the output of the converter as well as to suppress voltage droop caused by a step in the load current. The best capacitor for IVR is one that could be mounted as close to the load as possible.

3M developed a technology where capacitors can be placed in the package substrate directly underneath the load. The capacitors have much lower impedance to the load, which is causing the load current step. Figure 33 depicts these capacitors. They measure a mere 220 μm thick in a 1005 package at 1 μF capacitance [55].

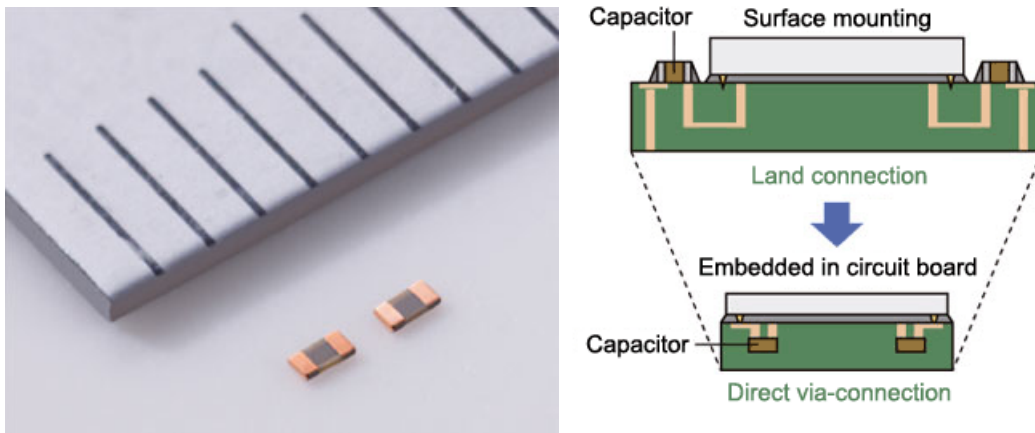


Figure 33: Tiny built-in capacitors (from [39])

Many manufacturers have driven down the path to miniature capacitors that can be embedded within a PCB or package substrate. Given that the capacitor can be mounted directly beneath the die, the ESR and ESL reduce dramatically.

4.2.1 Integrated Capacitor Options

The capacitor mounted inside the package substrate provides a generous improvement over having the capacitor on the motherboard. However, there is also an opportunity to have capacitance on the die itself.

Many options for integrated capacitors appear in today's processors. MOS capacitance typically provides 10 $\text{fF}/\mu\text{m}^2$ (10 nF/mm^2) or lower. Metal-metal capacitance is typically 5-10x less than the MOS capacitance density, or 1-2 nF/mm^2 .

Another up and coming technology that Intel published is the metal-insulator-metal capacitor in a 22 nm Tri-Gate process [5]. This capacitor, shown in Figure 34, stretches between the uppermost vias, making use of otherwise wasted space required for via separation without interrupting current paths.

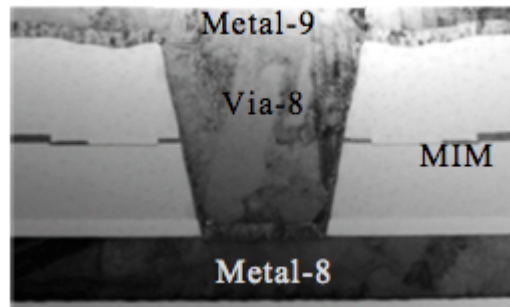


Figure 34: Intel 22 nm process MIM capacitor between M8 and M9 (from [5])

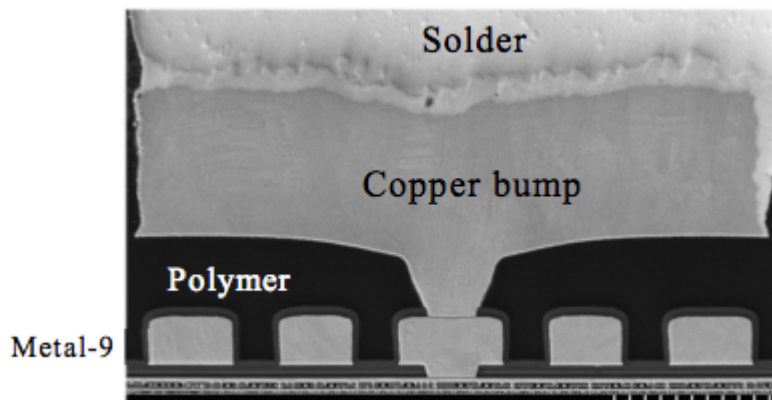


Figure 35: Full metal stack of Intel 22 nm Tri-Gate (from [5])

However, the on-die MIM cap yields only 20 nF/mm². While the density isn't as high as a discrete ceramic capacitor, the integrated capacitor has a higher bandwidth. That being said, the integrated capacitor requires Intel's proprietary processing steps.

4.3 Power Switches

Power FETs, like inductors and capacitors, may be fully integrated or discrete. One option may be to use the same transistors that the SOC makes use of. Another solution incorporates a separate die including the power FETs. A tiny die could solder to the package that includes power FETs of a different technology than the SOC itself, perhaps with improved performance for power delivery rather than computation.

Buck converters favor a higher input voltage to some degree. In contrast to a linear regulator, efficiency generally improves with higher voltage. A reduction of average input current into the buck for higher input voltages relaxes physical distribution requirements on the input rail. Infrastructure also demands that high voltage rails come into the chassis to reduce cabling requirements within the datacenter. For example, assume that V_{DDX} is the input voltage to the chip and V_{DDF} is the final logic supply powering the high performance core. The ratio N is defined as V_{DDX}/V_{DDF} . The average input current into the chip is reduced by N , playing the same game that the system architect plays by providing a monstrous 48 V input to the rack. Also, the switching regulator produces a higher current slew rate, which improves the dynamic load regulation.

Additionally, the duty cycle of the switching supply in continuous current mode, D , is roughly equal to N . During the charge phase of the switching supply, a PMOS power FET device steers current into the inductor from the upstream supply, assuming that there is no boosted gate NMOS device being implemented on the high side. The NMOS power FET steers current back into the inductor from ground during the synchronous phase. Increasing the input supply effectively means that the regulator spends more time with an NMOS conducting current rather than a PMOS. NMOS

devices, of course, have higher conductivity. So, the system benefits in efficiency by spending more time in the NMOS power FET phase.

However, CMOS logic voltages continue to trend downwards. A conflict exists between the needs of the logic transistors and the requirements of infrastructure distribution. The power converter provides the necessary transformation between infrastructure needs and high performance logic requirements. In order to balance the transformation both the infrastructure and the logic technology need to meet in the middle.

The logic technology, through device engineering, can be enhanced to support higher voltages. This can come in the form of drain extensions and thicker oxides. However, these can add cost and time to the technology development and take a backseat to the native logic devices. They suffer reduced performance compared to the native logic transistors. Foundries develop thick oxide transistors with a higher V_{MAX} than the core logic devices to ease the design of legacy I/O's operating at a higher voltage than the core logic. Technologies are not optimized for these lower speed I/O devices. Also, these "extra" devices cause extra development effort at the foundry as well as additional process steps that all add up to additional cost.

Circuit techniques can also raise the usable voltage range of the native logic FETs. Using stacked protection devices can extend the voltage range. However, the fundamental switch metrics still need to be considered.

4.3.1 Performance Metrics

Transistors serve many different functions and can be tailored for specific purposes. Power FETs basically operate as a switch in two states. An ideal power FET

possesses zero resistance while on, infinite resistance while off, and zero energy to transition between the two states. When comparing power FETs, the best way to level the playing field is to discuss them in terms of milliohm-nanocoulombs ($\text{m}\Omega\text{-nC}$). For example, a power transistor with a rating of $10 \text{ m}\Omega\text{-nC}$ might require 2 nC of gate charge to realize $5 \text{ m}\Omega$ from drain to source. The charge component implies that a pre-driver pushes (or pulls) the gate capacitance to a certain voltage level to reach the specified resistance from drain to source. A transistor with extremely low on resistance may require driving the gate to a prohibitively high level due to dynamic power. With a low $\text{m}\Omega\text{-nC}$ specification, there is some hope to having low loss while in the conduction state without requiring a large amount of dynamic power. The chart in Figure 36 depicts power FET performance for many discrete devices and a 45 nm Predictive Technology Model (PTM) [3].

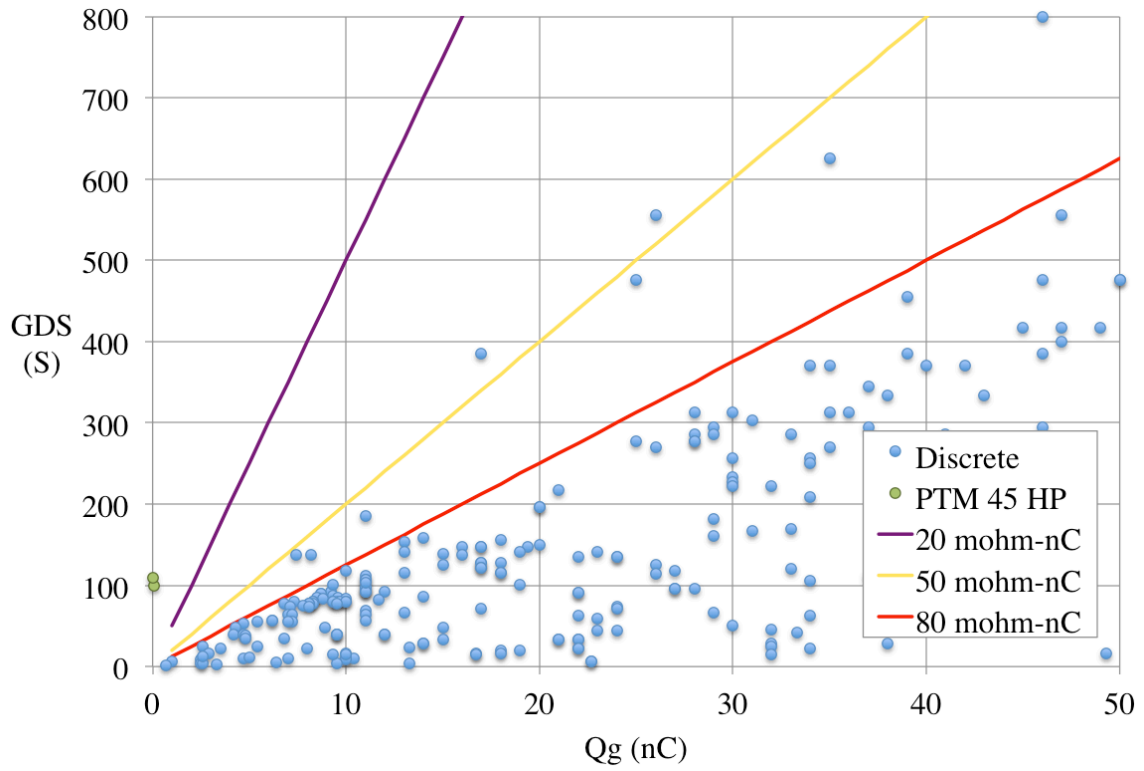


Figure 36: Power FET drain-source conductance versus Qgate

More detailed characterization data from HSpice of the 45 nm PTM model is depicted in Table 4.

Type	Parameter	Value	Unit
PMOS	V _{pnmon}	1.00	V
	W _{pmos}	1.00E-06	m
	L _{pmos}	4.50E-08	m
	m _{pmos}	4.00E+04	
	A _{pgatearea}	1.80E-09	mm ²
	Q _{p_{perum}}	1.50E-15	C/μm
	S _{pcondperum}	2.50E-03	S/μm
	C _{pgateperum}	1.50E-15	F/μm
	S _{pcondtotal}	1.00E+02	S
	R _{p_{total}}	1.00E-02	Ω
	Q _{p_{total}}	6.00E-11	C
	C _{p_{total}}	6.00E-11	F
NMOS	V _{nnom}	1.00	V
	W _{nmos}	1.00E-06	m
	L _{nmos}	4.50E-08	m
	m _{nmos}	2.50E+04	
	A _{pgatearea}	1.13E-09	mm ²
	Q _{nperum}	1.60E-15	C/μm
	S _{ncondperum}	4.37E-03	S/μm
	C _{ngateperum}	1.60E-15	F/μm
	S _{ncondtotal}	1.09E+02	S
	R _{n_{total}}	9.15E-03	Ω
	Q _{n_{total}}	4.00E-11	C
	C _{n_{total}}	4.00E-11	F

Table 4: Power FET characterization data from ASU 45 nm PTM model

In addition to these parameters, the amount of current pushed through the FET is thermally limited, as the power dissipation due to I^2R losses produces a substantial temperature rise. Given that these power FETs will live in close proximity to the main logic circuitry, the amount of heat generated by the power FETs has to be brought into the part's total power dissipation. In the past this power dissipation occurred out of the package, where it is now part of the packaged part. Thus, this “additional” power was already accounted for in the total system, but the location has changed.

Another key consideration is the maximum drain-to-source voltage (V_{DS}) supported. This voltage determines the DC or nominal input voltage to the regulator(s). A higher input voltage reduces the amount of current coming into the package. A higher input voltage also allows for a higher current slope in the inductor. Section 5.4 describes this in more detail.

The maximum gate-to-source voltage (V_{GS}) also comes into play from a reliability standpoint. When the current going through the inductor is interrupted, as is the case in buck converters, the inductor will drive the input voltage to a large negative level resulting from $V_{ind} = L di/dt$. This effect reduces as the transistor begins to conduct when the voltage drops sufficiently to a) forward bias the bulk-drain diode or b) cause a V_{GS} to develop via terminal reversal. This dip causes a V_{GS} excursion beyond the normal operation and may force the system VDD to be lower than the normal operating voltage of the power FET, which doesn't take full advantage of the technology.

Most off-the-shelf power FETs allow V_{DS} of 12, 24, or even 48 V. High performance CMOS logic transistors operate at less than 1 V and lower. However, a lower performance transistor with a thicker oxide and higher V_{MAX} usually coexists with the logic transistors to support legacy I/Os. As a result of the desire to have a large input

voltage and the voltage excursion, several circuit options can be utilized to support an input voltage higher than the thin oxide high-performance transistor V_{MAX} .

4.3.2 Single stacked Power FETs

A single transistor provides the simplest solution for the power FET. An NFET may take the place of P0 in an attempt to achieve higher performance in discrete designs. However, in a bulk CMOS application a non-VSS NFET body connection requires additional area and circuitry. The single stack solution is shown in Figure 37.

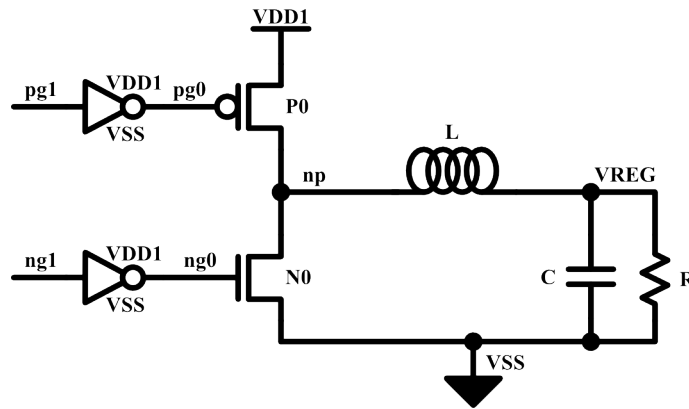


Figure 37: Schematic of single stacked power FETs.

The switching waveform for this simple case is shown in Figure 38.

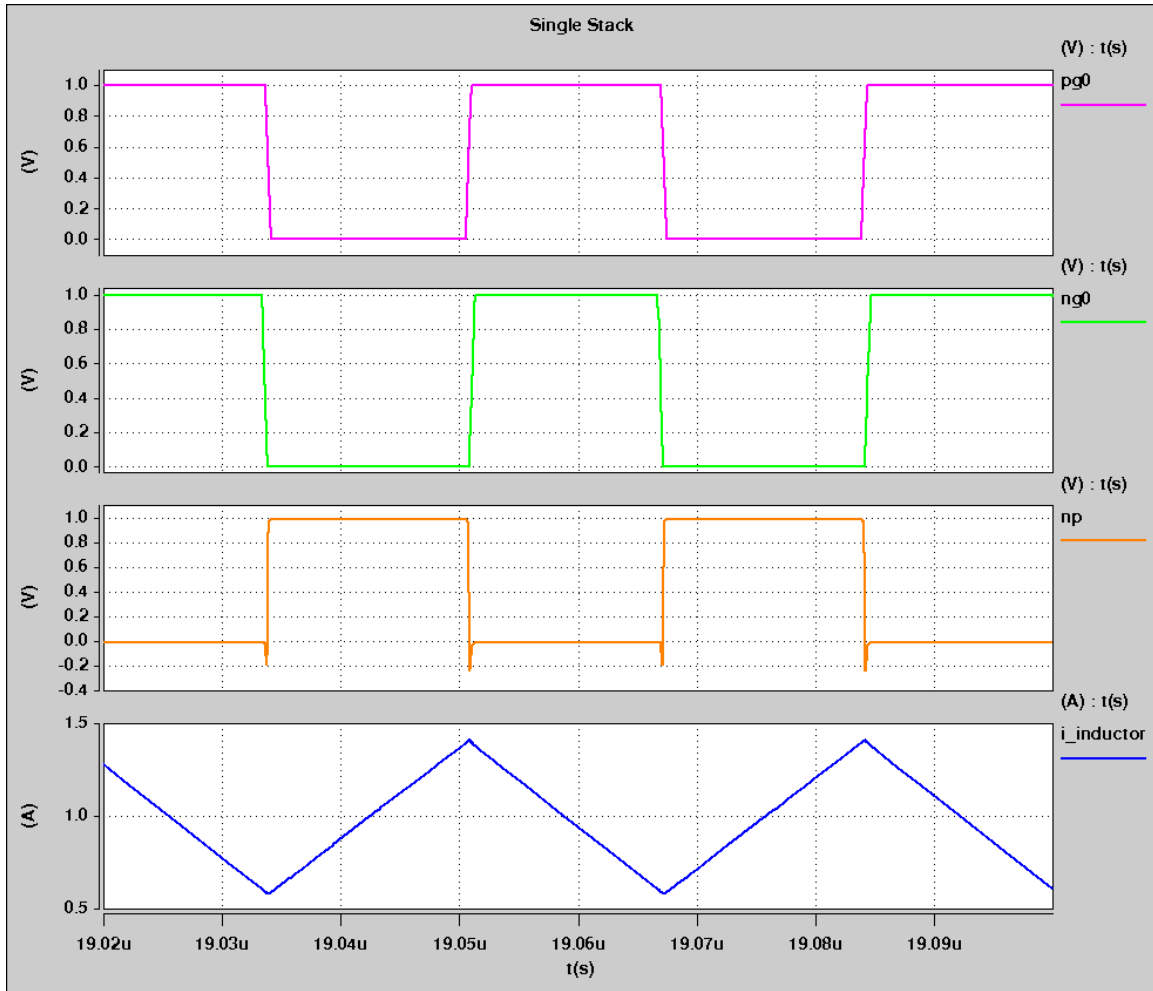


Figure 38: Switching waveform of single stacked power FET topology

As shown in Figure 38, the inputs $pg0$ and $ng0$ simply toggle at f_{sw} between VSS and VDD. A separate circuit adds non-overlap protection to prevent P0 and N0 from turning on at the same time. Note the excursion on node np during the transitions caused by interruption of the inductor current. The topology lacks level shifters, analog biases, and other complications that arise from double or triple stacked FETs.

The maximum VDD supported is easily determined from the technology max:

$$VDD1 = V_{max} \quad (11)$$

Conduction power in the power FETs results from the drain-source resistance of the power FETs multiplied by the average current. D denotes the duty ratio of PMOS and NMOS phases:

$$P_{cond} = I^2 [DR_{DSP} + (1 - D)R_{DSN}] \quad (12)$$

The dynamic power of the single stack bridge itself is driven mainly by the capacitance of the gates, the switching frequency, and the swing on the gate voltages.

$$P_{dyn} = (C_{gn} + C_{gp})V_{DD1}^2 f_{sw} \quad (13)$$

The drain capacitances also contribute to this dynamic power. However, a portion of that charge can be delivered to the load.

In terms of area, the single stack of power FETs provides a very dense physical unit. The tighter density also helps with parasitic resistance and capacitance. In many cases, the power FET density may need to decrease to achieve a lower power density or conform to the Electro-Migration (EM) limits of the controlled collapse chip connector (C4) bumps.

4.3.3 Double Stacked Power FETs

If the transistors do not support the desired input voltage, a stacked structure provides higher input supply as shown in Figure 39.

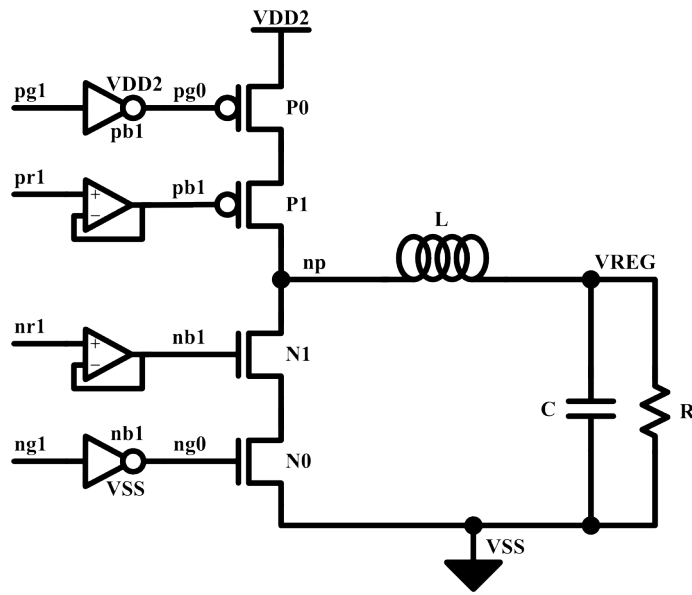


Figure 39: Schematic of double stacked power FETs.

This structure is common for high voltage I/O's. Amplifiers bias the gates of P1 and N1 to “protect” P0 and N0 from the full VDD level.

The waveforms for the double stack power FETs are depicted in Figure 40.

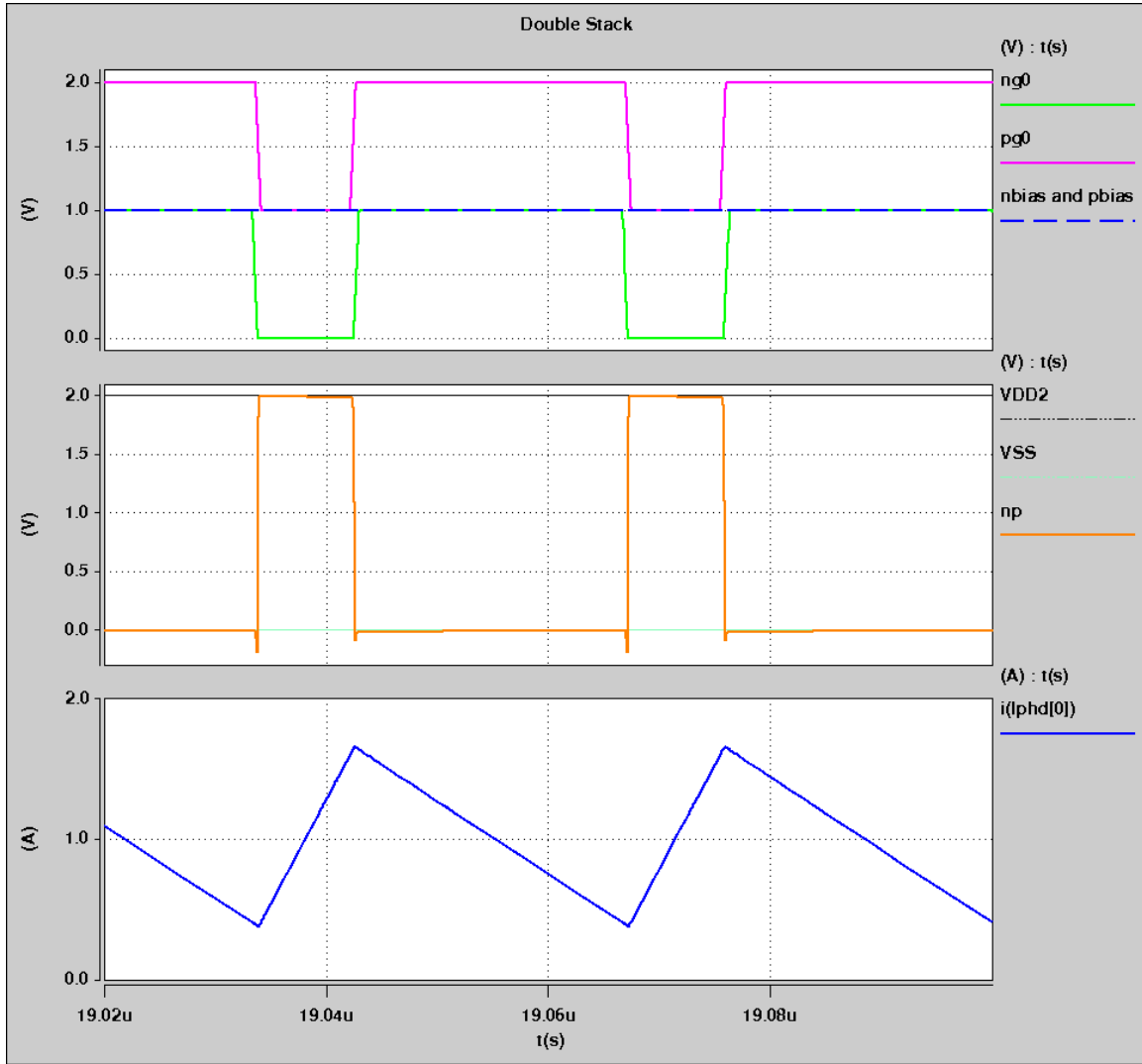


Figure 40: Switching waveform of double stack power FETs.

In theory, the maximum input voltage that this structure can support reaches:

$$VDD2 = 2 \cdot V_{max} \quad (14)$$

In this case biases pb1 and nb1 can combine and be set to $VDD2/2 = V_{max}$. If $VDD2$ is lowered, as is often required in practice, then nb1 may increase and p1 may decrease to maintain $V_{gs} = V_{max}$, making the most of the available technology.

Note that the area of the double stack, assuming the same overall resistance target as the single stack, will have grown to 4x the single stack size. This comes from the doubling of the widths and the fact that you need two devices in the first place.

However, the switching gate capacitance only grows by 2x when compared to the single stack topology.

The pre-driver input for P0 now needs a level shifter to protect the devices in the input inverter. This level shifter doesn't consume a substantial amount of power, but can complicate the overlap timing.

The bias levels nb1 and pb1 must be driven from a low impedance amplifier or low pass filter from a ratio of the input supply. This circuit may draw significant power in order to reach a sufficiently low impedance to operate as intended. The parasitic Miller capacitance results in significant coupling from the output node back onto the bias signal, decreasing the effectiveness of the biased protection device.

4.3.4 Triple Stacked Power FETs

If the double stack power FET topology still does not provide a high enough input voltage, then a triple stack power FET topology may suffice, but with significantly increased complexity, area, and power. This topology, documented in 2004, implements an active protection bias [31].

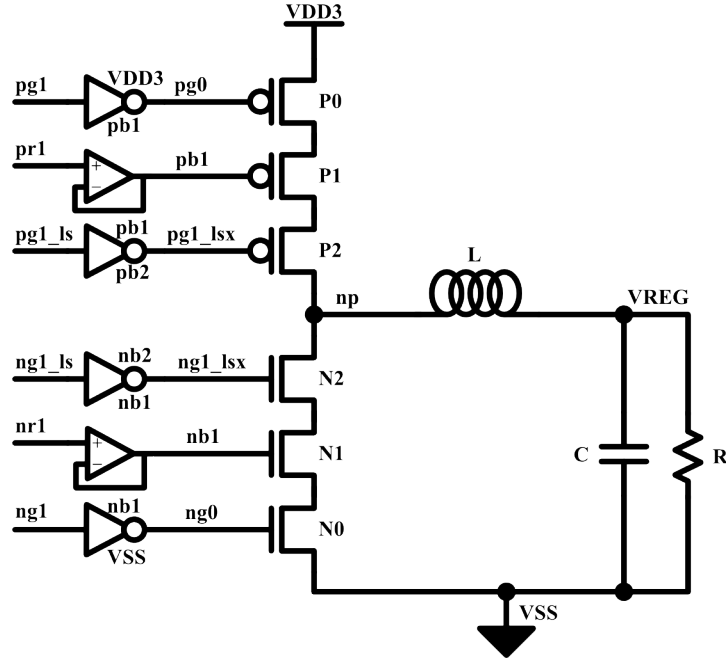


Figure 41: Schematic of triple stacked power FETs for higher input supply.

The outer FETs, N0 and P0, still operate as in the single and double stack case. P0's gate is driven up to VDD3 and down to pb1. N0's gate is driven from VSS up to nb1. P1 and N1 bias from pb1 and nb1, operating exactly as in the double stack scenario.

The inner devices, P2 and N2, bring on the challenge. They operate as “active cascodes.” When P0 is off, the drain of P2 sees VSS, which places the full output voltage across the PFET stack. So, pg1_ls biases to $VDD3 - 2 \cdot V_{max}$, which protects P1. P1 then protects P0 just as in the double stack case. When P0 turns on, the drain and source of P2 rise to VDD3. The gate must increase to $VDD3 - 1 \cdot V_{max}$ to avoid an oxide overstress issue. Therefore, the gate of P2 rises while the gate of P0 falls to adjust the biasing every period. The same biasing scheme occurs for N2, N1, and N0 with respect to VSS.

Figure 42 depicts the waveforms for the triple stack power FET topology. Note that the current in the inductor ramps up at a faster rate, but the ramp down remains at the same rate regardless of the input voltage.

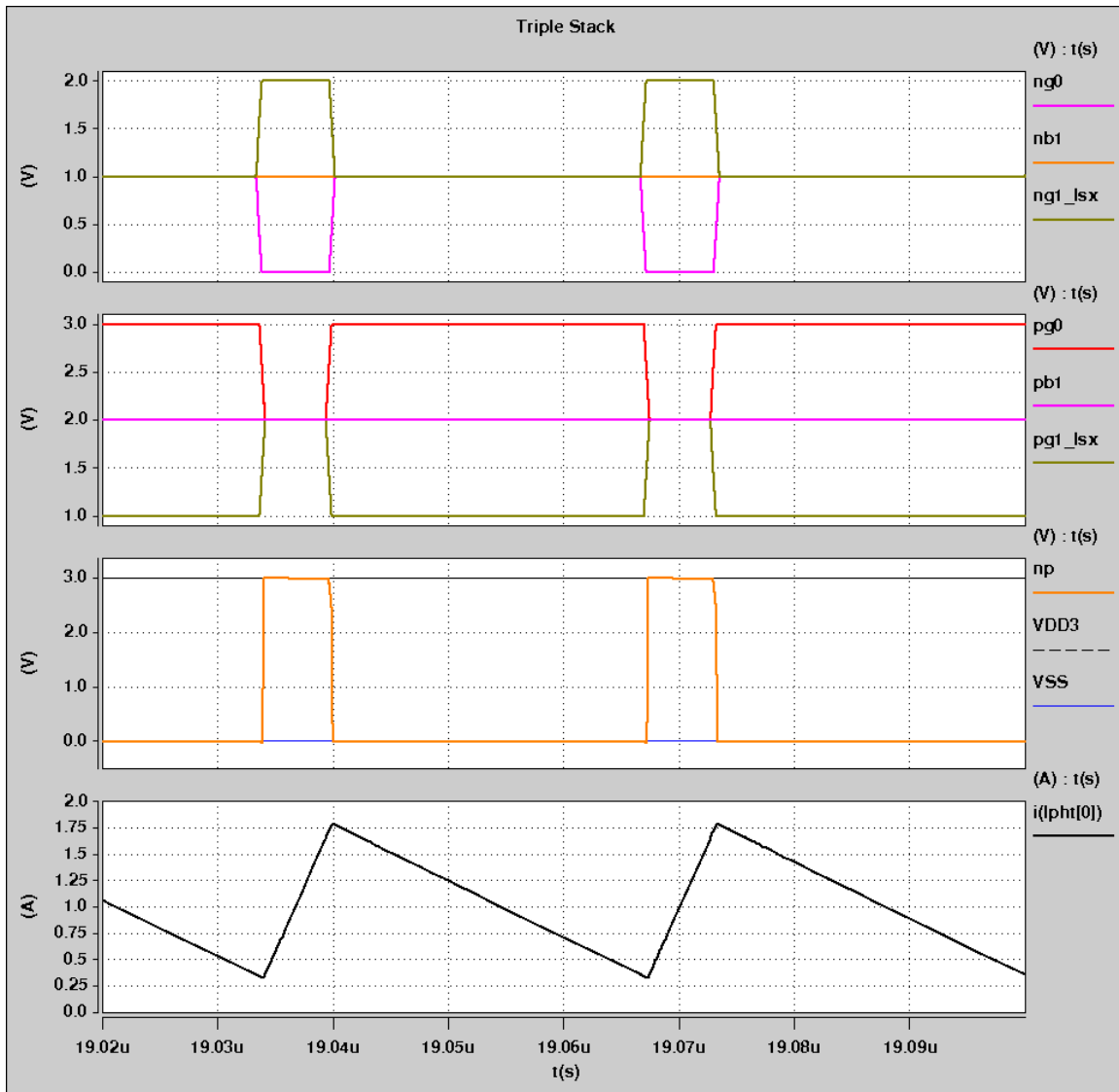


Figure 42: Switching waveform of the triple stack power FET topology

Assuming a comparable impedance target as for the single stack case, 9x area is required, a huge increase.

When compared to the single stack with equivalent R_{DS} , $2 \cdot 3 = 6$ additional gate capacitances toggles when accounting for gate dynamic power. The gates are 3x wider, and the active protection-bias-signals switch in addition to the primary gates (2x).

The design requires four analog voltage generators (pb1, pb2, nb1, and nb2), all with strict output impedance requirements. Three level shifters accommodate the active protection biases and the pre-driver for pg0.

The active protection bias extends the transistor technology to high voltage, but can be dangerous because of the timing of the active protection bias relative to the drain node movement. Ensuring that gate oxide overstress does not affect the device is challenging. Additionally, the body voltages must be driven in the protection FETs to prevent gate-body overstress effects.

Nonetheless, support for a high input voltage comes with many advantages. The input current through the package decreases. Inductor current slew rates increase, allowing for faster transient slew rates.

4.3.5 Summary of Stacking Capability

In summary, higher input voltages can be supported, but the circuit complexity increases dramatically. The following table describes the approximate trend for single, double, and triple stacked devices as a guide.

Metric	Single	Double	Triple
V_{in}	$1 \cdot V_{max}$	$2 \cdot V_{max}$	$3 \cdot V_{max}$
Area	1x	4x	9x
C_{gatesw}	1x	2x	6x
mΩ-nC	1x	2x	6x
m_1	$(V_{max} - V_{out})/L$	$(2V_{max} - V_{out})/L$	$(3V_{max} - V_{out})/L$
m_2	$-V_{out}/L$	$-V_{out}/L$	$-V_{out}/L$
Amplifiers	0	1-2	3-4
Level Shifters	0	1	3

Table 5: Stacked power FET performance summary

Metrics m_1 and m_2 specify the current slew rates in a buck converter during charge and synchronous rectification phases, respectively [3]. During the charge phase, the inductor integrates the difference between input and output supply, which is essentially a constant. During the synchronous rectification phase, the inductor integrates the output supply only.

$$m_1 = \frac{V_{IN} - V_{OUT}}{L} \quad (15)$$

$$m_2 = \frac{-V_{OUT}}{L} \quad (16)$$

This chapter provides information about key elements of the IVR solution. The availability of these elements and their performance feeds into the system design stage of development as well as the circuit level design.

4.4 Summary

This chapter provides information about the fundamental power converter elements from an integrated regulator perspective.

Inductors store energy in the form of current. A recent surge in handheld devices has driven development in miniature inductors. Generally an integrated inductor does not perform as well as a discrete one because of the process constraints. However, discrete inductors pose form factor challenges. A suitable inductor for integrated regulation has a high current capacity, low DC resistance, high series resonant frequency, and moderate inductance. Inductance for integrated regulation lies in the 10-100 nH range.

Recently developed capacitors allow for embedding within the package core, thereby reducing the parasitic elements between the capacitor and the load that it decouples. Metal-insulator-metal capacitors may also be implemented in the upper metal layers. These capacitors possess lower capacitance per unit area, but significantly lower parasitic resistance and inductance.

For ease of integration the simplest way to implement power FETs is to use thick oxide I/O transistors. A high input supply proves advantageous for increased current slew rate and lower average input current. Transistor technologies continue to push the operational voltage lower. Stacking transistors allows the architect to take advantage of higher input supplies given a low voltage transistor.

These fundamental elements give rise to an overall architecture for an integrated power supply solution.

5 ARCHITECTURE

This chapter discusses system-level architecture including the processor load, requirements for the on-die IVR, and external components from Chapter 4. The architectural phase of design leads into lower-level circuit design and analysis.

5.1 IVR Practicality

Integrating the voltage regulator onto an SOC presents a major upheaval and brings many risks. To make matters worse, the choice must be made during the early design-plan phase because of the extent of the upheaval. For some systems, IVR removes external components to shrink the system. Other systems see IVR as an avenue to power reduction. High performance microprocessors have reached a thermal limit where IVR provides a path to increased performance. The reduction in external components comes second to power reduction.

The architect first considers the primary elements of Chapter 4: inductor, capacitor, and power FET. The processor logic load, or the main logic IP, provides additional constraints. These feed into the overall IVR architecture. The architecture, in turn, provides guidance to the detailed circuit design. These steps also feed back into previous steps. Then circuit design, or schematics, leads to physical layout. During each step the designer must look forward to consider the roadblocks ahead carefully.

Many system aspects push regulators to integration. Inductors shrink dramatically. Capacitance densities continue to rise. Semiconductor technologies marched on to allow consumer electronics to become more portable. Laptops became thinner. In 2003 a high performance CPU core ran at 1.4 GHz in a desktop tower. Today dual 1.3 GHz cores run in palm-sized smartphones. The trend towards miniaturization has

pushed the market for microscopic passive elements, opening the door for the migration of power regulation closer to the CPU core.

5.2 Load Definition

A high performance microprocessor named Tuolumne provides a test case for the integrated regulator. The microprocessor targets the server segment for high-density rack systems. The servers provide a compute solution for a large database engine company. The software program follows a simple pattern. The processor sleeps, wakes up to service a database query, and then returns to sleep. The server's performance relies on waking up quickly, operating at a high performance mode, and then quickly returning to sleep.

5.2.1 Physical Constraints

Tuolumne provides four CPU cores, a dual core graphics engine, DDR interface, and several high-speed serial links. The general-purpose core spans an area of roughly 30 mm². The graphics processor makes use of two units at 50 mm² per unit. IOs and miscellaneous occupy the remaining 50 mm². The total comes to 270 mm². Dies usually stretch one dimension slightly longer than the other. A width of 18 mm and height of 15 mm provide a reasonable footprint.

C4 bumps connect the die to the package. These array in a close packed hexagonal (CPH) structure, depicted in Figure 43, to maximize the bump density and therefore the amount of metal that can make the connection from die to package.

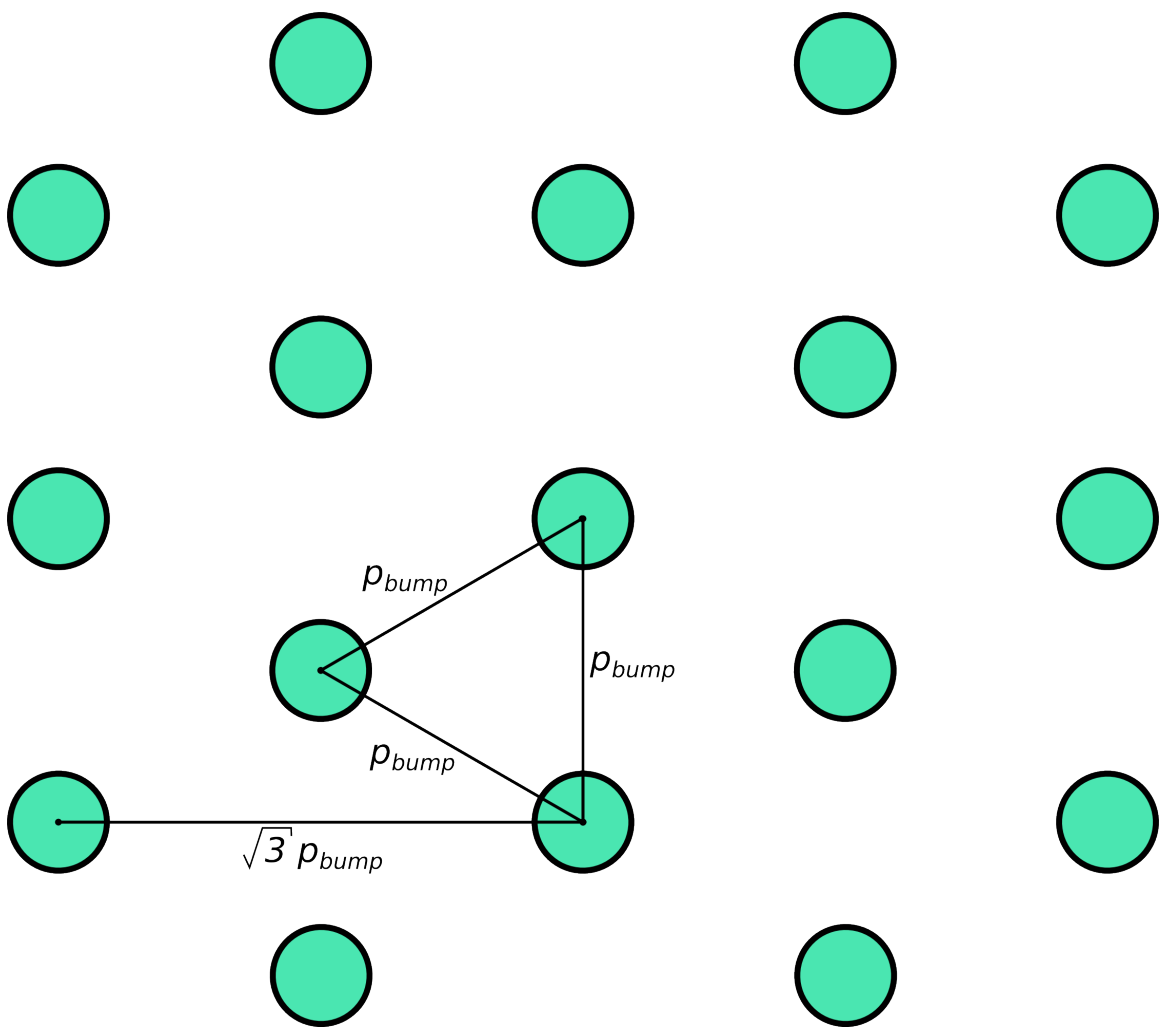


Figure 43: Close-packed hexagonal bump spacing.

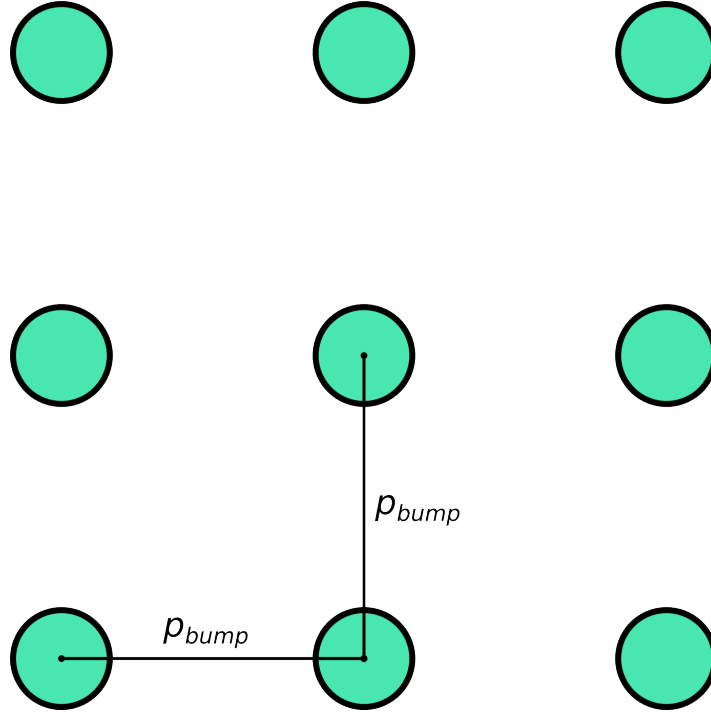


Figure 44: Simple grid structure

The simple grid structure, shown in Figure 44, does not make optimal utilization of the provided area. In one square of $p_{bump} \times p_{bump}$, only one bump exists.

$$\text{Cubic Density} = \frac{1 \text{ bump}}{p_{bump}^2} \quad (17)$$

The structure in Figure 43 provides a 15% higher bump density as compared to Figure 44. Consider the equilateral triangle of sides p_{bump} . This triangle encloses $1/6^{\text{th}}$ of 3 bumps. This and the area of the equilateral triangle reveals the CPH bump density:

$$\text{CPH Density} = \frac{\frac{1}{2} \text{ bump}}{\frac{1}{2} p_{bump}^2 \sin(60^\circ)} = \frac{1 \text{ bump}}{\frac{\sqrt{3}}{2} p_{bump}^2} = 1.15 \cdot \text{Cubic Density} \quad (18)$$

The CPH structure has a colorful history. Sir Walter Raleigh tasked his mathematician, Thomas Harriot, to come up with an efficient way to count cannonballs

stacked on the decks of his ships around 1587. Harriot later influenced Johannes Kepler, who proposed the famous Kepler conjecture. Kepler stated in 1611 that under this structure “the packing will be the tightest possible.” This conjecture was listed in David Hilbert’s list of twenty-three unsolved mathematical problems of 1900. The conjecture remained without a formal proof until Thomas Hales, a Texan, provided a computer aided proof in 1998 [23].

Given a bump pitch, p_{bump} , and the area of the die, A_{die} , one may determine approximately how many bumps the processor has provided by taking the die area divided by the area one bump occupies in a CPH crystal structure. High performance parts maintain a bump pitch as low as 150 μm .

$$N_{bump} = \frac{A_{die}}{\frac{\sqrt{3}}{2} p_{bump}^2} \quad (19)$$

A package provides a transformation from the uppermost die metal wiring to the connecting motherboard PCB routing. Pin pitch, p_{pin} , ultimately bounds the package area along with the number of pins, N_{pin} . Strangely, pins typically do not array in the hexagonal shape. They array in simple square grids instead. Therefore, the approximate package size, $A_{package}$, can be determined:

$$A_{package} = N_{pin} r_{pin}^2 \quad (20)$$

Modest package pin pitches lie in the 1.2 mm range. Tuolumne possesses 1,021 pins, in line with modern server package capabilities. This yields a package area of 1,470 mm^2 . The form factor favors a non-square rectangular package shape similar to the die. A 42 mm by 35 mm package fits well. Both package and die have an aspect ratio of 1.2.

Table 6 describes Tuolumne’s physical dimensions of the large IP as well as the package dimensions.

Specification	Value	Unit
NCPU	4	x
NGPU	2	x
AREACPU	30	mm ²
AREAGPU	50	mm ²
AREAOTHER	50	mm ²
DIEAREA	270	mm ²
DIEASPECT	1.2	x
XDIE	18	mm
YDIE	15	mm
PBUMP	150	μm
NBUMPS	13856	x
NPINS	1021	x
PPIN	1.20	mm
PKGAREA	1470	mm ²
PKGASPECT	1.20	x
XPKG	42	mm
YPKG	35	mm

Table 6: Tuolumne physical specifications

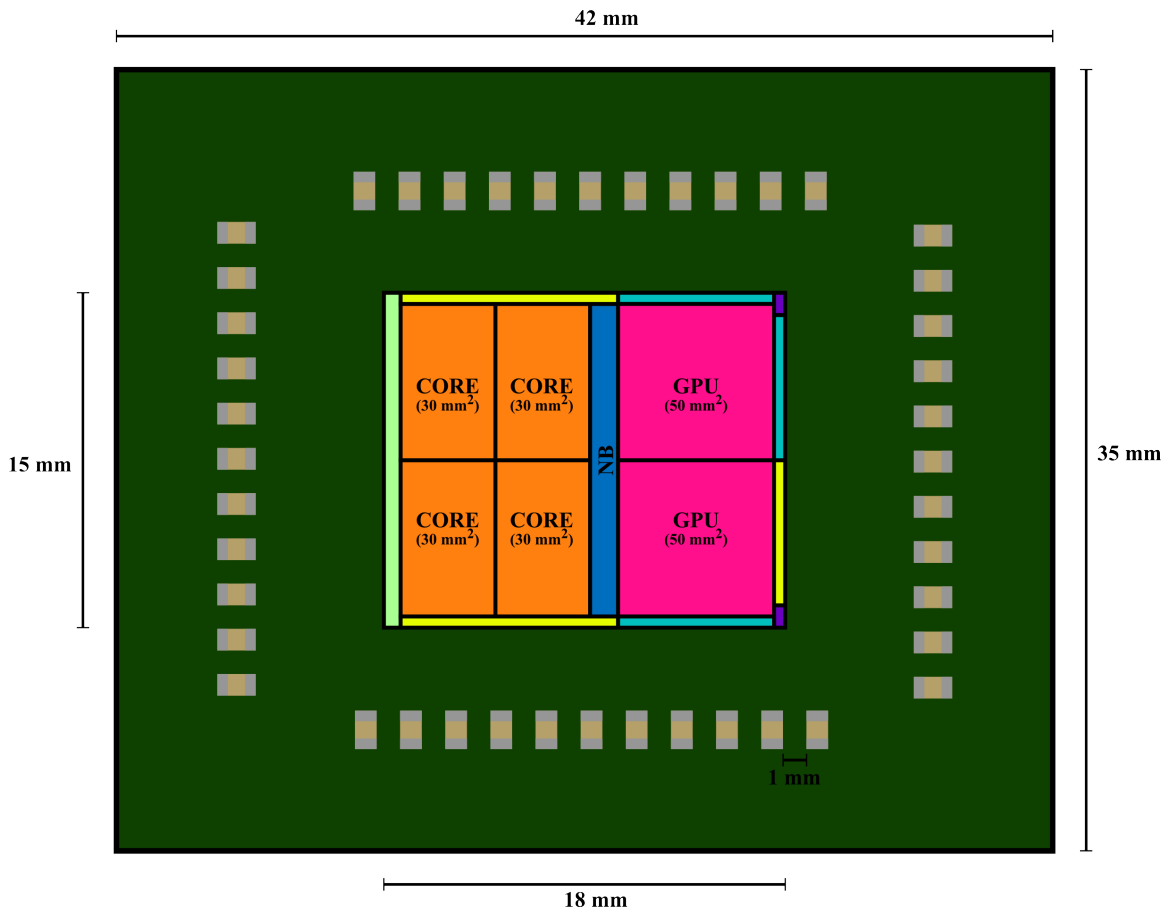


Figure 45: Scale drawing of the Tuolumne high performance server chip

Tuolumne will serve as the example case for IVR. The following definitions allow for the progression of Tuolumne from an external voltage regulator (EVR) to an IVR. The die may grow slightly from the original Tuolumne specifications to allow for additional power FETs and active circuitry associated with the IVR. The overall package allows minimal change because the motherboard form factor cannot change drastically.

5.2.2 *Electrical Constraints*

Having outlined the physical constraints of the load, the next logical step is to move on to the electrical requirements.

Tuolumne possesses a maximum power state where the voltage regulator supplies the maximum transistor voltage at maximum frequency. This combination produces the highest power density the silicon supports.

Standard CMOS technology has a sustained power density limit, W_{max} , of around 2 W/mm^2 without the use of exotic phase-change cooling technologies. Beyond that limit, irreparable damage occurs. This limit has effectively capped frequency scaling of high performance CPUs since cooling technology has remained basically the same (large metal heat sink with a fan). Heat dissipation now caps the frequency [25]. This arises largely from the inability to scale V_t , a function of kT/q [15]. While this power density limit provides a theoretical maximum power density limit for the load, Tuolumne follows a slightly lower power density for thermal margin.

The maximum power density, W_{max} , and maximum transistor voltage, V_{DDmax} , result in a maximum current density, I_{max} in terms of A/mm^2 . Optimization studies by Nose and Sakurai published in 2000 proved that optimal low power and high-speed sizing results in 30% of the total power being attributed to leakage [42]. This provides a floor for the current at a given voltage.

Table 7 enumerates the electrical constraints of the Tuolumne loads based on these limits.

Specification	Value	Unit
PWRDENSMAX	1.30	W/mm ²
Leak Ratio	0.3	x
Area	30	mm ²
Instances	4	x
VDD	1.25	V
FSW	3000	MHz
CSWEFF	5.8	nF
CSWDENS	194.1	pF/mm ²
IDDMAX	31.2	A
IDDDYN	21.8	A
IDDSTATIC	9.4	A
IDDMAXDENS	1.0	A/mm ²
IDDDYNDENS	0.7	A/mm ²
IDDSTATICDENS	0.3	A/mm ²
IMAXTOTAL	124.8	A
PMAX	39.00	W
PDYN	27.30	W
PSTATIC	11.70	W
PMAXDENS	1.300	W/mm ²
PDYNDENS	0.910	W/mm ²
PSTATICDENS	0.010	W/mm ²
PTOTAL	156.00	W
RMAX	133.5	m Ω
RMIN	40.1	m Ω

Table 7: Tuolumne CPU electrical load constraints

These specifications are comparable to high performance Intel blade server processors in production from HP. Similar formulations apply to the GPU load, which presents a lower power density and switching frequency.

Specification	Value	Unit
PWRDENSMAX	0.30	W/mm ²
Leak Ratio	0.3	x
Area	50	mm ²
Instances	2	x
VDD	1.20	V
FSW	600	MHz
CSWEFF	12.2	nF
CSWDENS	243.1	pF/mm ²
IDDMAX	12.5	A
IDDDYN	8.8	A
IDDSTATIC	3.8	A
IDDMAXDENS	0.3	A/mm ²
IDDDYNDENS	0.2	A/mm ²
IDDSTATICDENS	0.1	A/mm ²
IMAXTOTAL	25.0	A
PMAX	15.00	W
PDYN	10.50	W
PSTATIC	4.50	W
PMAXDENS	0.300	W/mm ²
PDYNDENS	0.210	W/mm ²
PSTATICDENS	0.002	W/mm ²
PTOTAL	30.00	W
RMAX	320.0	mΩ
RMIN	96.0	mΩ

Table 8: Tuolumne GPU electrical load constraints

Note that the GPU power density achieves only around a quarter of the CPU power density, yielding a lower overall power for the GPU.

5.3 Phase-Centric Designs

With load requirements fully quantified in the previous chapter along with the elements of Chapter 4, opportunities for power conversion on the package present themselves. Section 4.1 goes into great length describing inductor technology available and how that is fundamental to IVR. Section 5.2 enumerates the current requirements of the regulator based on high performance CMOS design.

Given the difficulties in producing an on-die power inductor in a foundry technology, the inductor is selected from discrete devices on sale currently from inductor manufacturers.

On-package inductors impose a practical limit on the number of inductors that can be placed on the package due to routing breakout of I/Os. Metal traversing to and from the inductor will require a lot of routing resources in its own right. That being said, the package area consumes a much larger area than the die. This is a result of the packages purpose in the first place. The package spatially translates PCB geometries to IC geometries. This transformation occupies a large amount of space when you consider a package with 1000+ pins residing underneath it. This allows a lot of room to place inductors for IVR on top of the package itself.

The Tuolumne package and die dimensions dictate the allowable area for inductor placement. Suppose that an array of inductors form a rectangular perimeter around the die, allowing equidistant space to the package edge and die edge as in Figure 45. The allowable clearance for the inductor to the die and package edge would be:

$$x_{clearance} = \frac{(x_p - x_d)}{4} - \frac{x_L}{2} \quad (21)$$

The approximate linear distance of the perimeter can also be determined:

$$L_{dist} = x_{pkg} + x_{die} + y_{pkg} + y_{die} \quad (22)$$

5.3.1 Coilcraft 0908SQ-27N_L Phase

From the commercially available inductors in Section 4.1.2, a suitable inductor should have a high current density and high inductance without too much DC Resistance (DCR). The Coilcraft 0908SQ-27N_L air core wire wound inductor measures 27 nH and 700 mA/mm², yet only 10 mΩ of DCR. The inductor's dimensions are 2.97 mm by 2.13 mm. With the long axis pointing away from the die the linear perimeter calculated in (22) and 0.25 mm spacing allows for 34 inductors. Each inductor carries a maximum current of 4.4 amps, which would provide 150 amps of current. This barely meets the total capacity required for both the CPU and GPU complexes in Tuolumne. However, pushing the inductors out slightly towards die edge provides 0.5 mm spacing and 40 inductors. Figure 46 depicts this placement option.

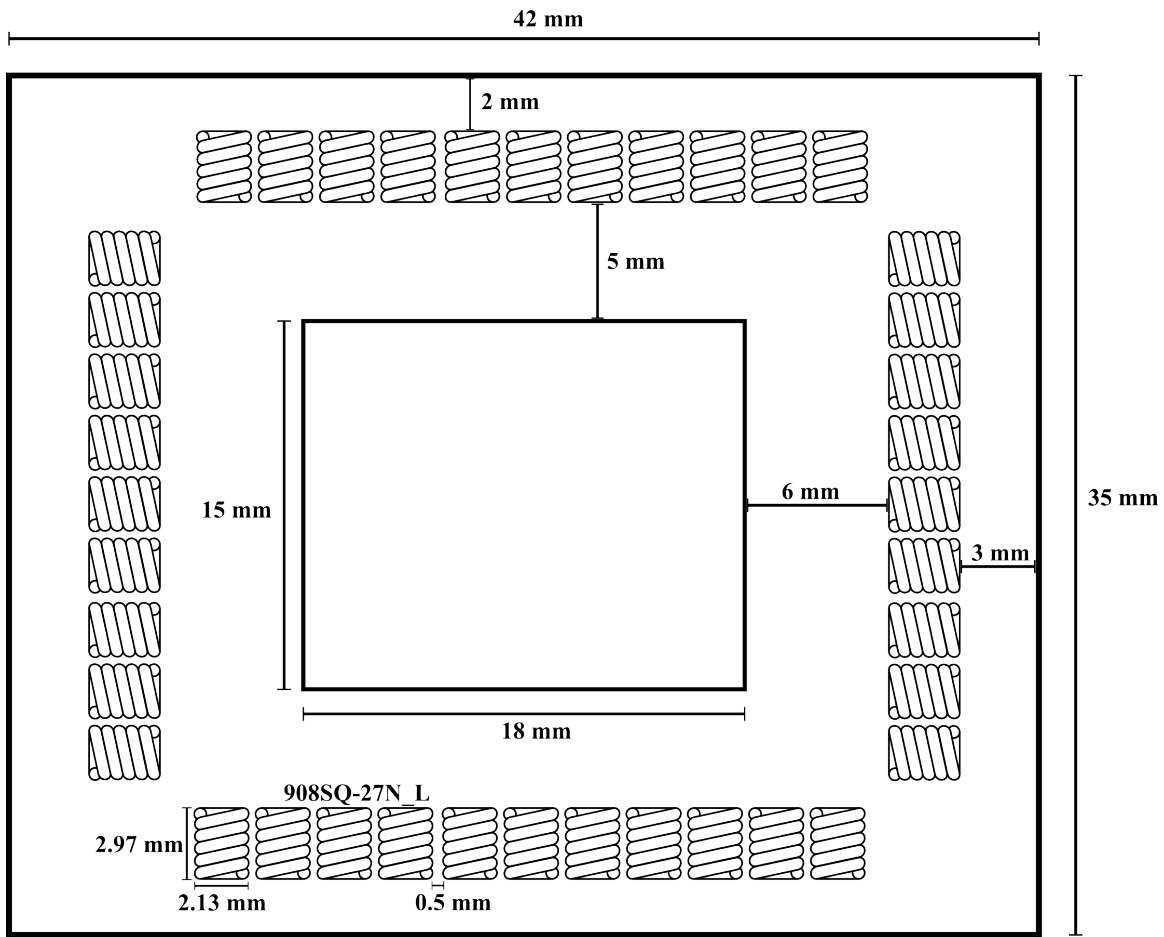


Figure 46: Scale drawing of 40 Coilcraft 908SQ-27N_L 4.4 A inductor placements

The package allocates 50% of the package area beneath the die to embedded capacitors. The TDK CGB2A1JB1C474K033BC embedded capacitor provides suitable bulk capacitance for the IVR solution. This 0402 capacitor, with a conservative 0.5 mm keep-out in each direction, occupies 1.5 mm² of package area. With this density, the 30 mm² CPU load allows placement of 10 TDK embedded capacitors directly beneath the load. At 2.2 μ F per 0402 MLCC capacitor the total comes to 22 μ F with an ESR of the parallel capacitors of 1.1 m Ω .

Power FETs on the active die toggle the input to the inductor. A stack of two 45 nm 1 V power FETs optimistically favors a 2 V input voltage source. With inductive kick and power supply ripple, the maximum V_{GS} will exceed 2 V to some degree. This assumes that no other high voltage power FET device would be available, forcing a 2-high stack configuration. Otherwise the input supply would be too close to the output supply, which penalizes the performance of the converter.

The normalized drain-source resistances of the PMOS and NMOS devices are $400 \Omega\text{-}\mu\text{m}$ and $229 \Omega\text{-}\mu\text{m}$, respectively. Targeting $10 \text{ m}\Omega$, a 2-high stack of 80,000 $1 \mu\text{m}$ PFETs each comprise the switch and protection device. Similarly for the NFET, 50,000 FETs form the switch, and 50,000 FETs form the protection device. The area utilization of the power FETs (gate area/FET area) must be driven below 10% to keep the power density of the power FETs themselves below 2 W/mm^2 . The high performance transistors occupy around 6% of the load area due to the power density constraint. The actual FETs themselves occupy only 0.3%. Conduction loss in the power FETs, on average, totals 147 mW.

Dynamic power from the power FET gate capacitance detracts from the efficiency of the solution. From the PTM data characterized in Table 4, the charge per micron of gate width of the PMOS and NMOS devices at 45 nm gate length are fixed at $1.5 \text{ fC}/\mu\text{m}$ and $1.6 \text{ fC}/\mu\text{m}$, respectively. The gate voltage of the protection FETs is fixed at a bias voltage, so the dynamic power predominately comes from the gate of the switching FETs. A loss of 100 mW is allocated to the low impedance source, which drives the gate of the N and P protection devices. 80,000 μm of PFET and 50,000 μm of NFET both toggle to 1 V at the switching frequency of 35 MHz. So, the approximate dynamic power of the NFET and PFET switches comes to 3 mW and 4 mW, respectively. Add to these

the power to buffer the gate signal, approximately twice the dynamic power of the gate itself to be conservative. In all the dynamic power comes to 21 mW.

Paths to shuttle current from the active die to the inductor and back constitute a substantial amount of loss to the overall delivery. Most of this comes in the form of copper on the package with a resistivity of $1.68\text{e-}8 \text{ } \Omega\text{-m}$. The routing to the inductor is assumed to be, on average, 1.5 mm wide. This is a conservative estimate for the worst-case copper routing to the inductor and back. The inductor itself has a width of 2.13 mm, but the path narrows down near the die to escape the C4 array. The return trace can be made much wider since the output of several inductors can be shorted to form a multi-phase supply whereas the input to each inductor must be unique. With $\frac{1}{4}$ -oz copper in the trace and an approximate round trip length of 10 mm, the per-phase copper resistivity estimate reaches 13 m Ω per phase. This contributes 192 mW power loss per phase.

Another source of loss comes into play. The controller consumes some amount of parasitic power to operate amplifiers, comparators, logic, etc. The assumption is made that the controller burns 25 mA of fixed current from the 2 V supply in addition to a proportional current that is 0.5% of the load. In many cases a current signal within the analog controller may be required that tracks the output current. The proportional term accounts for this power loss.

In all the per-phase power efficiency of this architecture and configuration crests 87% at around 3.5 W output power per phase, or 23 A total output current for the 8 phases at 1.2 V.

Each phase of the converter, including inductor and capacitor, occupies 6.4 mm² where the load occupies 3.8 mm². The converter delivers power at a density of 0.434 W/mm². In other words, 1 mm² of power supply circuitry can deliver 434 mW.

5.3.2 Coilcraft 0807SQ-11N_L Solution

A smaller Coilcraft inductor, the 0807SQ-11N_L supports a 2.7 A current rating, but only 11 nH. This inductor provides a smaller footprint at 1.55 mm x 1.83 mm. Tuolumne package physical specifications allow 48 of these inductors for a total of 130 A capability.

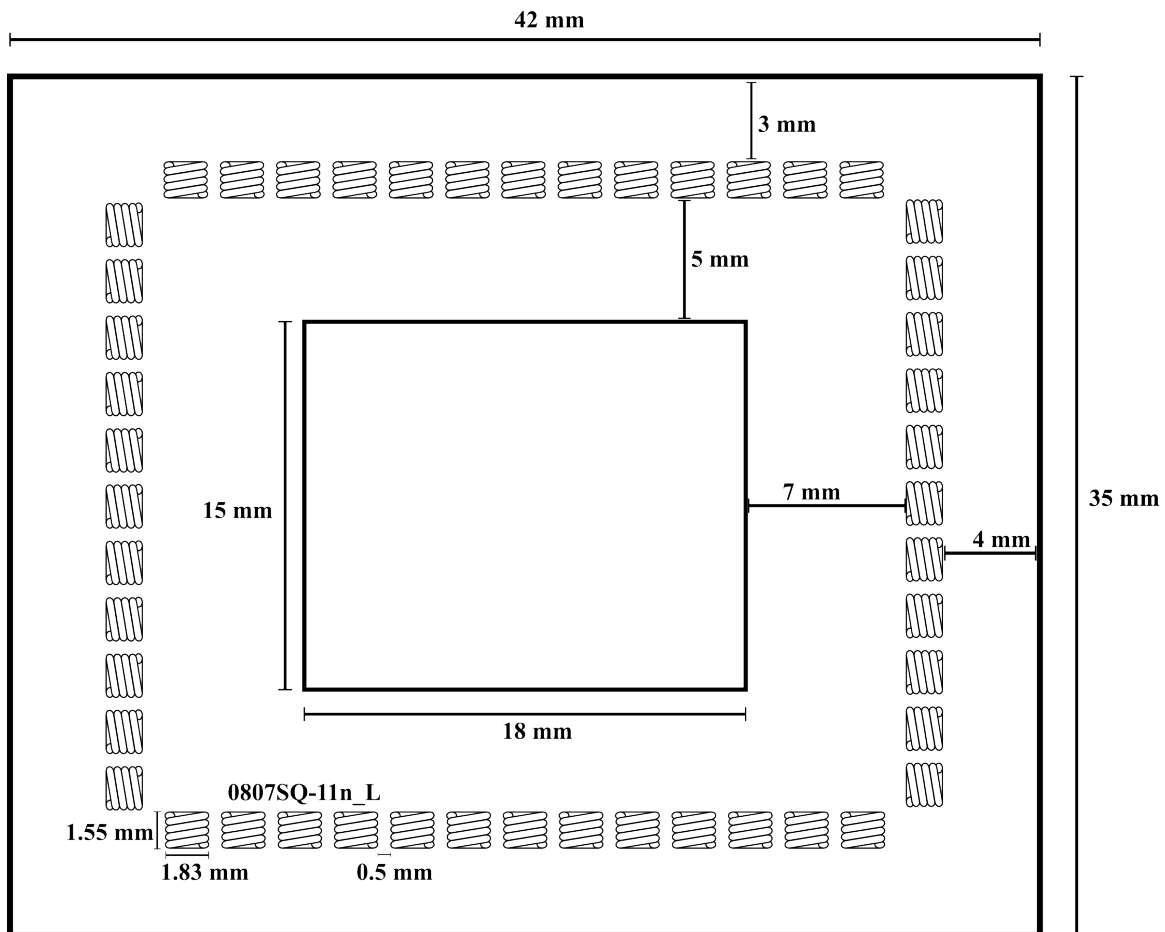


Figure 47: Scale drawing of Coilcraft 0807SQ-11N_L 2.7 A inductor placements

Both of these inductor solutions provide a massive amount of current on the package that is comparable to the conventional regulator on the motherboard. Provided

that the currents remain the same, the first solution maps directly. If the load current can be reduced the lower current solution could work as well.

Note that this solution, in theory, could produce a separate rail for each inductor, yielding upwards of 48 independent rails for the chip. A conventional system would likely provide a single core supply for all of the cores and a single supply for the graphics. The next logical step, with this in mind, is to design a single converter phase.

The efficiency of the 0807SQ-11N solution is comparable to that of the 0908SQ-27N. The 0807SQ-11N provides a finer granularity and lower height restriction.

5.4 Voltage Droop

One of the primary goals of IVR is to improve the response of the regulator to a step increase in output current at the load. If the load current suddenly increases, which happens quite frequently in high performance microprocessors, the output voltage will droop. The amount of droop results from the passive devices in the voltage regulator design more so than the controller. By the time the analog controller responds, the damage has already been done. Careful engineering in the passives can reduce the voltage droop as a result of the load current increase.

A conventional voltage regulator benefits from a large array of bulk decoupling capacitors that reduce the droop in addition to high frequency package capacitors. As the power routing gets closer to the processor, the capacitors become smaller and affect high frequency noise better. However, IVR pushes the inductor quite close to the load and provides more individual supplies, removing the large low frequency capacitors. This results in less total capacitance on each rail. The bandwidth of the IVR is pushed higher by using a smaller inductor, reducing the total capacitance required. Package-based IVR

also pushes the capacitor to be placed directly under the load circuit to free up space for the inductor. This reduces parasitic routing effects that the load sees to the capacitor.

Consider the buck regulator circuit of Figure 48. The power FETs, P0 and N0, drive inductor L in series and capacitor C to VSS. At $t = 0$, the switch closes, causing a sharp increase in output current from the regulator. Current through the inductor in a buck configuration ramps up and down at $m_1 = (VDD - VREG)/L$ and $m_2 = -VREG/L$.

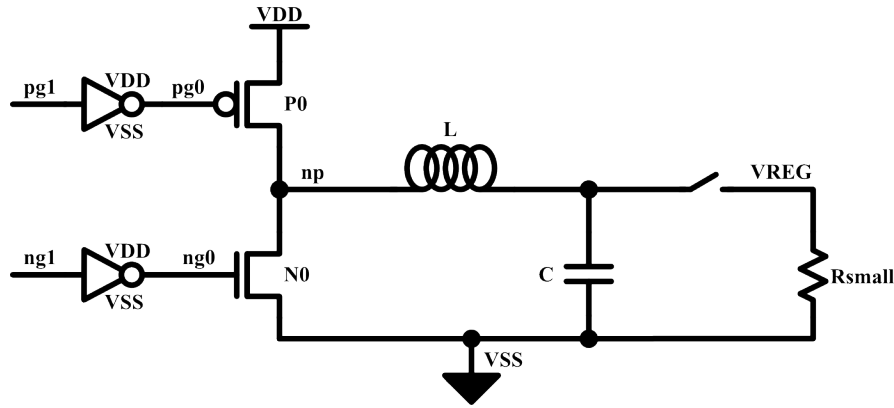


Figure 48: Schematic of regulator with load step

For this analysis the assumption will be that the switch closes at the beginning of the m_1 cycle, where the inductor current begins its ramp up. This is optimistic for a single phase, but realistic for multi-phase designs. With this in mind the inductor model consists of a current source with initial value i_0 and a current ramp rate of $m_1 = (VDD - VREG)/L$. The voltage across the inductor is considered to be much larger than ripple and the droop, thereby keeping $VDD - VREG$ as a constant in the model. Figure 49 displays the model with the power FETs and inductor modeled as a current source.

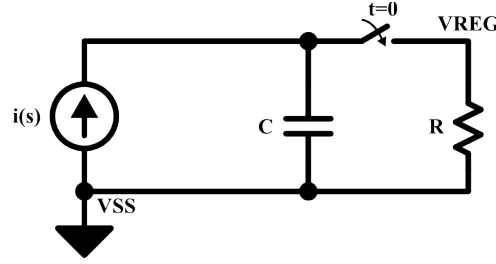


Figure 49: Reduced model for derivation of maximum V_{droop} as a result of instantaneous current step

Therefore, the input to the system, $i_{in}(s)$, is:

$$i_{in}(s) = \frac{i_0}{s} + \frac{m_1}{s^2} \quad (23)$$

The load, Z_L , is the parallel combination of the resistor and capacitor where $\omega_p = (RC)^{-1}$. When the current step occurs, the capacitor immediately supplies charge, as the current through the inductor cannot change instantaneously.

$$Z_L(s) = \frac{1}{C} \left(\frac{1}{s + \omega_p} \right) \quad (24)$$

Under this formulation the output voltage may be constructed simply as the input current multiplied by the load impedance.

$$v_{out}(s) = i_{in}(s) \cdot Z_L(s) \quad (25)$$

$$v_{out}(s) = \left(\frac{i_0}{s} + \frac{m_1}{s^2} \right) \frac{1}{C} \left(\frac{1}{s + \omega_p} \right) \quad (26)$$

$$v_{out}(s) = \left(\frac{i_0}{C} \right) \frac{1}{s(s + \omega_p)} + \left(\frac{m_1}{C} \right) \frac{1}{s^2(s + \omega_p)} \quad (27)$$

In order to solve the inverse Laplace transform, the two terms undergo partial fraction expansion. First the $\frac{i_0}{C}$ term.

$$\frac{1}{s(s+\omega_p)} = \frac{1}{\omega_p} \left(\frac{1}{s} - \frac{1}{s+\omega_p} \right) \quad (28)$$

Then the $\frac{m_1}{C}$ term can be determined.

$$\frac{1}{s^2(s+\omega_p)} = \frac{1}{\omega_p^2} \left(\frac{\omega_p}{s^2} - \frac{1}{s} + \frac{1}{s+\omega_p} \right) \quad (29)$$

Now the equation for $v_{out}(s)$ is in a form suitable for the inverse Laplace transform.

$$v_{out}(s) = \left(\frac{i_o}{C\omega_p} \right) \left(\frac{1}{s} - \frac{1}{s+\omega_p} \right) + \left(\frac{m_1}{C\omega_p^2} \right) \left(\frac{\omega_p}{s^2} - \frac{1}{s} + \frac{1}{s+\omega_p} \right) \quad (30)$$

$$\mathcal{L}^{-1}\{v_{out}(s)\} = v_{out}(t) = \frac{i_o}{C\omega_p} [1 - e^{-\omega_p t}] + \frac{m_1}{C\omega_p^2} [\omega_p t - 1 + e^{-\omega_p t}] \quad (31)$$

With regrouping and the substitution of ω_p the output voltage is described as the sum of exponential and linear terms.

$$v_{out}(t) = (i_o R - m_1 R^2 C)(1 - e^{-t/(RC)}) + m_1 R \cdot t \quad (32)$$

This provides some very interesting observations. The linear term, $m_1 R \cdot t$, is simply the inductor current ramp producing a positive linear change in voltage. The first exponential term, $(i_o R)(1 - e^{-t/(RC)})$, represents the initial energy on the capacitor decaying with the RC time constant. The second exponential, $(-m_1 R^2 C)(1 - e^{-t/(RC)})$, represents the fact that the inductor delivers current to the capacitor and the load.

It is interesting to note that the derivatives of $m_1 R \cdot t$ and $(-m_1 R^2 C)(1 - e^{-t/(RC)})$ are equal for small t .

However, this does not consider the fact that the capacitor begins with an initial voltage, VREG, to begin with. With this initial condition applied to the voltage on the capacitor at $t = 0$, $v_{out}(0)$, the droop as a function of time becomes:

$$v_{out}(t) = (i_0R - m_1R^2C)(1 - e^{-t/(RC)}) + m_1R \cdot t + v_{out}(0)e^{-t/(RC)} \quad (33)$$

This may also be rewritten in terms of raw exponent, linear, and constant terms.

$$v_{out}(t) = (v_{out}(0) + m_1R^2C - i_0R_L)e^{-t/(RC)} + m_1R \cdot t + i_0R_L - m_1R^2C \quad (34)$$

Figure 50 depicts $v_{out}(t)$ for example case where the current ramped from 0.5 to 10 amps through a 20 nH inductor with a capacitance of 10 μ F. The voltage across the inductor is 1.8 V.

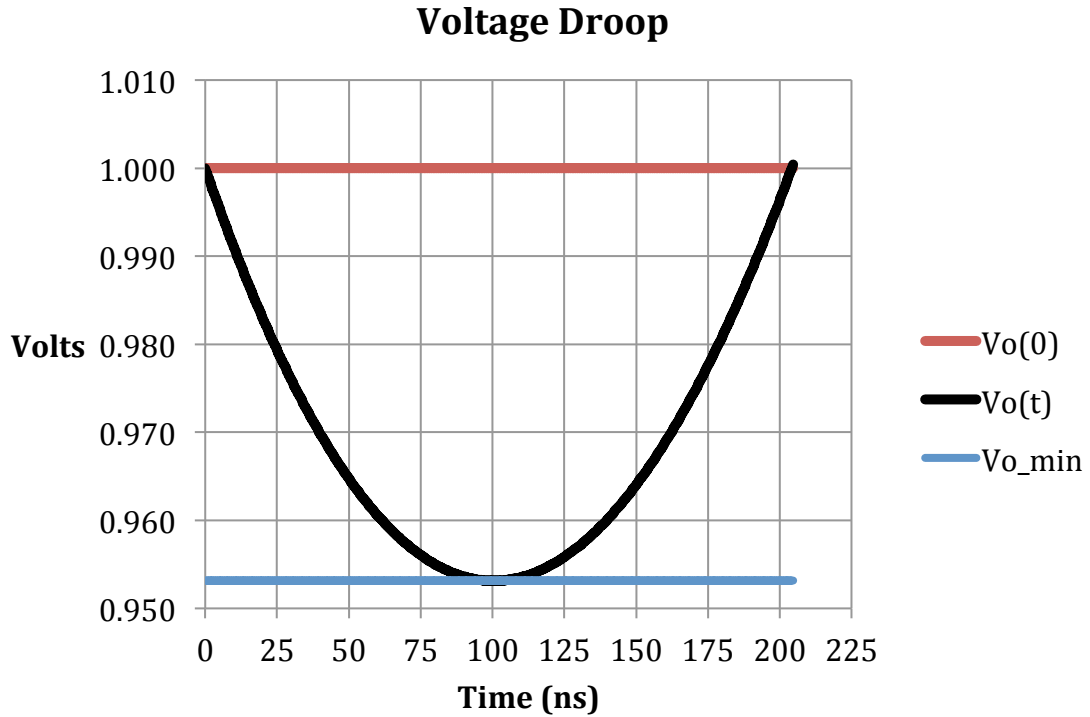


Figure 50: Voltage droop according to inverse Laplace model.

The voltage dips to 953 mV at 100 ns. At 200 ns the ramping inductor current replenishes the potential energy on the capacitor and therefore the original output voltage.

Now that a description of the output droop voltage has been formulated, the next logical question will be to determine the minimum droop level before the linear inductor current matches the resistor current such that the voltage droop stops.

This implies finding the derivative of $v_{out}(t)$.

$$\frac{dv_{out}(t)}{dt} = \frac{-1}{RC} (v_{out}(0) + m_1 R^2 C - i_0 R_L) e^{-t/(RC)} + m_1 R \quad (35)$$

Setting the derivative to zero allows finding t_{min} , the time at which the minimum of the function occurs.

$$\frac{-1}{RC} (v_{out}(0) + m_1 R^2 C - i_0 R_L) e^{-t_{min}/(RC)} + m_1 R = 0 \quad (36)$$

$$t_{min} = -RC \cdot \ln \left(\frac{m_1 R^2 C}{m_1 R^2 C + v_{out}(0) - i_0 R_L} \right) \quad (37)$$

Substituting the time for the minimum droop gives $v_{out}(t_{min})$ as:

$$v_{out}(t_{min}) = i_0 R_L - m_1 R^2 C \cdot \ln \left(\frac{m_1 R^2 C}{m_1 R^2 C + v_{out}(0) - i_0 R_L} \right) \quad (38)$$

With this solution the voltage droop, Δv_{out} , can be expressed in terms of the passive elements alone.

$$\Delta v_{out} = v_{out}(0) - i_0 R_L + m_1 R^2 C \cdot \ln \left(\frac{m_1 R^2 C}{m_1 R^2 C + v_{out}(0) - i_0 R_L} \right) \quad (39)$$

This solution provides a very good estimate of the best-case voltage droop in a regulator without parasitic inductance or capacitor ESR.

This analysis assumes that the current step occurs while the energy in the inductor is increasing, which may be considered optimistic. The timing of the load step can occur during the synchronous rectification phase while the inductor current is decreasing. In a

single-phase design with synchronized switching (non-hysteretic control), the regulator will wait until the next rising edge of clock before it can begin ramping the inductor current again to the load, relying on the capacitor to hold the voltage until the inductor can provide an increase in current. However, additional phases reduce this worst-case wait time dramatically as phases are added. In multi-phase regulators only a very small fraction of time exists when there would be no positive inductor current ramp from one of the phases. One option to alleviate this concern in single-phase designs would be to use a hysteretic controller, whereby the timing is fully asynchronous. In that case the controller can, within its bandwidth limitations, flip over to the PMOS phase when the load step occurs.

Additionally, capacitors and the connections to them possess parasitic resistance and inductance, limiting the capacitors ability to respond to the current step. As soon as the output load pulls a large load transient, the voltage will immediately droop to $\Delta I/RESR$ as a result of the ESR. However, this resistive portion can add directly to the solution for the passive network response derived already.

5.5 Load Line and Automatic Voltage Positioning

Intel introduced the Pentium® 4 microprocessor line in November 2000. At this time microprocessor currents reached levels that required new regulator techniques. Load currents began to ramp from near nothing to 50-60 amps in a few clock cycles, basically instantaneous from the motherboard voltage regulator's perspective. This di/dt in concert with the power delivery network produces tremendous droops on the power supply. In order to withstand such a tremendous di/dt , the concept of Automatic Voltage Positioning (AVP) [62] was created.

Conventional buck regulator control loops provide an integral term with high gain that will drive the output voltage back to the reference level regardless of the output load condition. In other words, the supply has a DC output resistance of 0 Ω . However, during the seemingly short time that the output voltage droops by Δv_{out} , hundreds of microprocessor clock cycles occur. Consider a droop of 100 ns duration, very fast in terms of voltage regulators. A 4 GHz microprocessor core undergoes 400 clock cycles in 100 ns. To avoid breaking speed paths the full-chip timing must comprehend this lowest voltage experienced from the voltage droop. Therefore, there is no incentive for the regulator to drive the voltage back up to the nominal level. AVP specifies that the regulator output impedance, at DC, be tightly set to R_{LL} , the load line resistance. Thus, the output voltage is regulated as

$$V_{out} = V_{ref} - R_{LL} \cdot I_{out} \quad (40)$$

Figure 51 depicts the same regulator with AVP on and off. Initially the current is effectively zero. The positive current step (0 A to 10 A) in the output causes the voltage to droop in both cases down to 953 mV. The AVP On case sustains the voltage near 956 mV while the AVP Off case returns the voltage to 1000 mV. The time between 200 ns and 600 ns, the AVP On case burns less power. Perhaps more importantly, note that as the current drops (10 A to 0 A), the AVP Off case overshoots to nearly 1050 mV, whereas the AVP On case overshoots back to the original 1000 mV operating condition. The AVP On case sustaining a 44 mV drop at 10 amps is termed a 4.4 m Ω load line resistance.

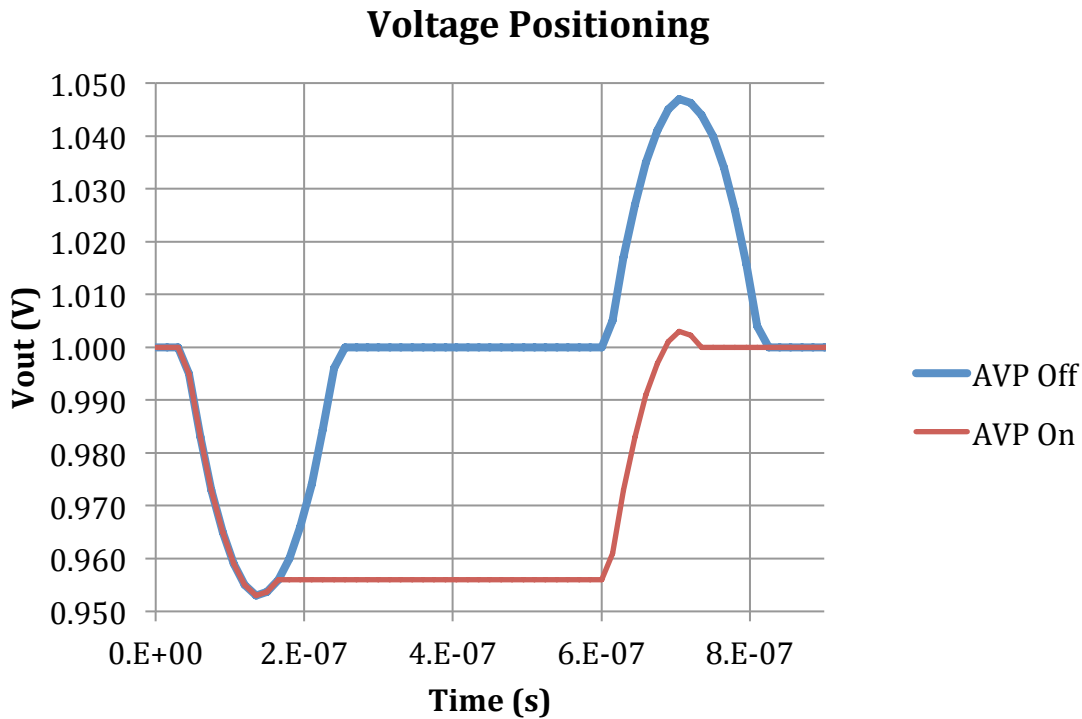


Figure 51: Positive and negative current steps with and without AVP.

The peak voltage that the power supply reaches affects the part's reliability. If the voltage overshoot is high enough and frequent enough, the gate oxide can degrade, potentially breaking a speed path or causing a logic fault altogether. So, the processor designers are squeezed from the low side to meet timing and the high side for reliability. AVP gives relief to these.

Ideally, the power delivery voltage droop matches the load line resistance exactly. If the power delivery network improves, the load line resistance can reduce. Figure 52 depicts a typical high performance processor's load line of only 2.0 mΩ. Most external VRMs allow programmable load line impedances. In an ideal scenario the load line resistance causes the output to sit exactly at the minimum voltage droop level. However, in some cases setting it slightly higher may trade reliability for timing margin.

Additionally, the load line implementation of AVP is a linear approximation to the droop. Voltage droop as a result of transients is not linear, which leaves some performance on the table.

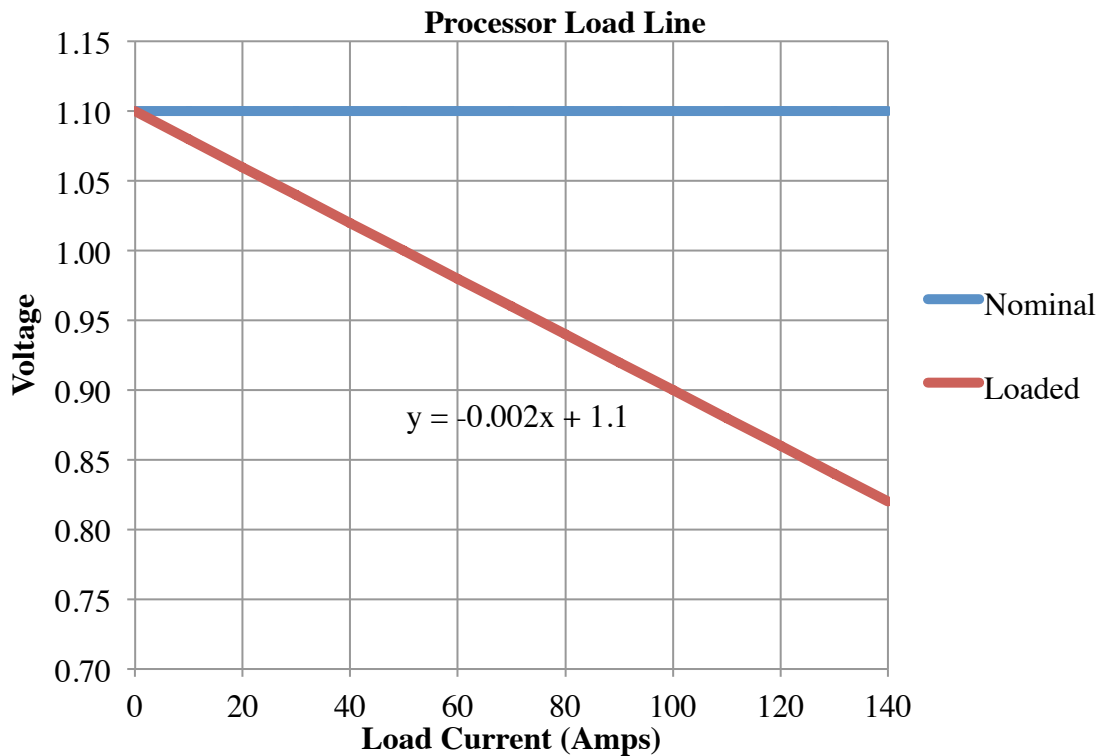


Figure 52: Graph of a 2.0 mΩ load line for a high performance microprocessor

This curve in particular speaks to an opportunity for improvement achievable only by IVR. In a high current state, such as 120 amps, the processor input voltage purposefully drops 240 mV below the VID. Another way to think about this would be to say that the processor must be overdriven by 240 mV in order to pass speed paths at max current. This is a direct result of the tremendous current ganged into one supply. With IVR the supplies can be split into many rails, each with a significantly lower current.

All of these considerations feed into the control and block-level circuit design. Architectural studies here show that IVR will fit, using commercially available inductors and capacitors. This also puts on display some of the high level features required of the controller.

5.6 Efficiency

As mentioned in section 5.3.1, the Coilcraft 27 nH solution surpasses 87% peak power efficiency at 3.5 W of power output per phase, or 28 W of total power output for the whole regulator. This corresponds to around 23 A at 1.2 V. The 13% of power lost is attributed to conduction loss, dynamic power, and control overhead, illustrated in Figure 53.

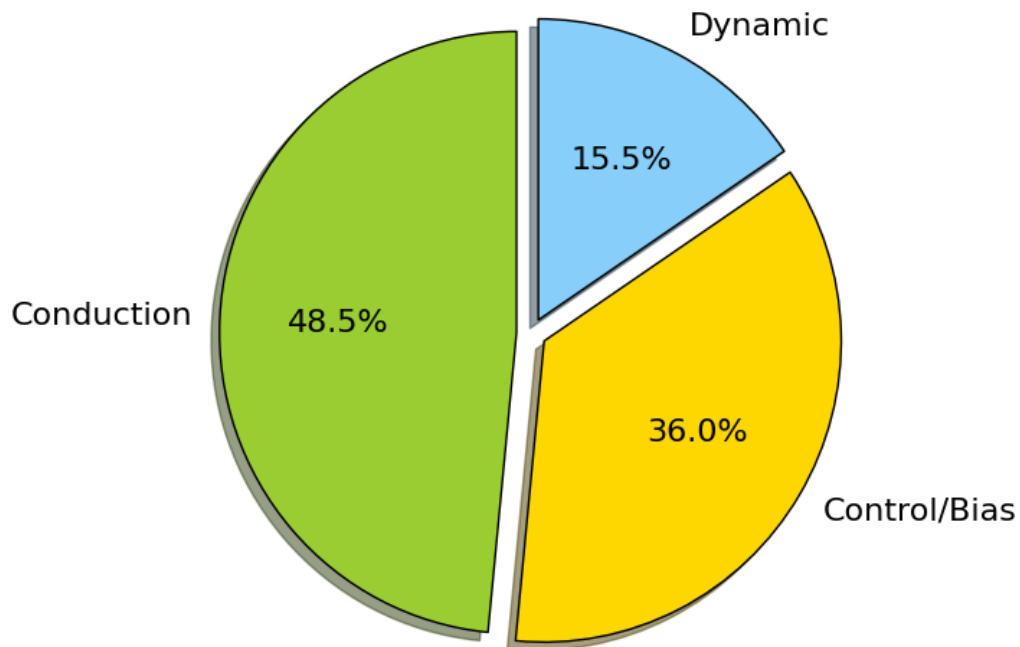


Figure 53: Breakdown of power loss in 27 nH phase

The total regulator efficiency versus output current is drawn in Figure 54.

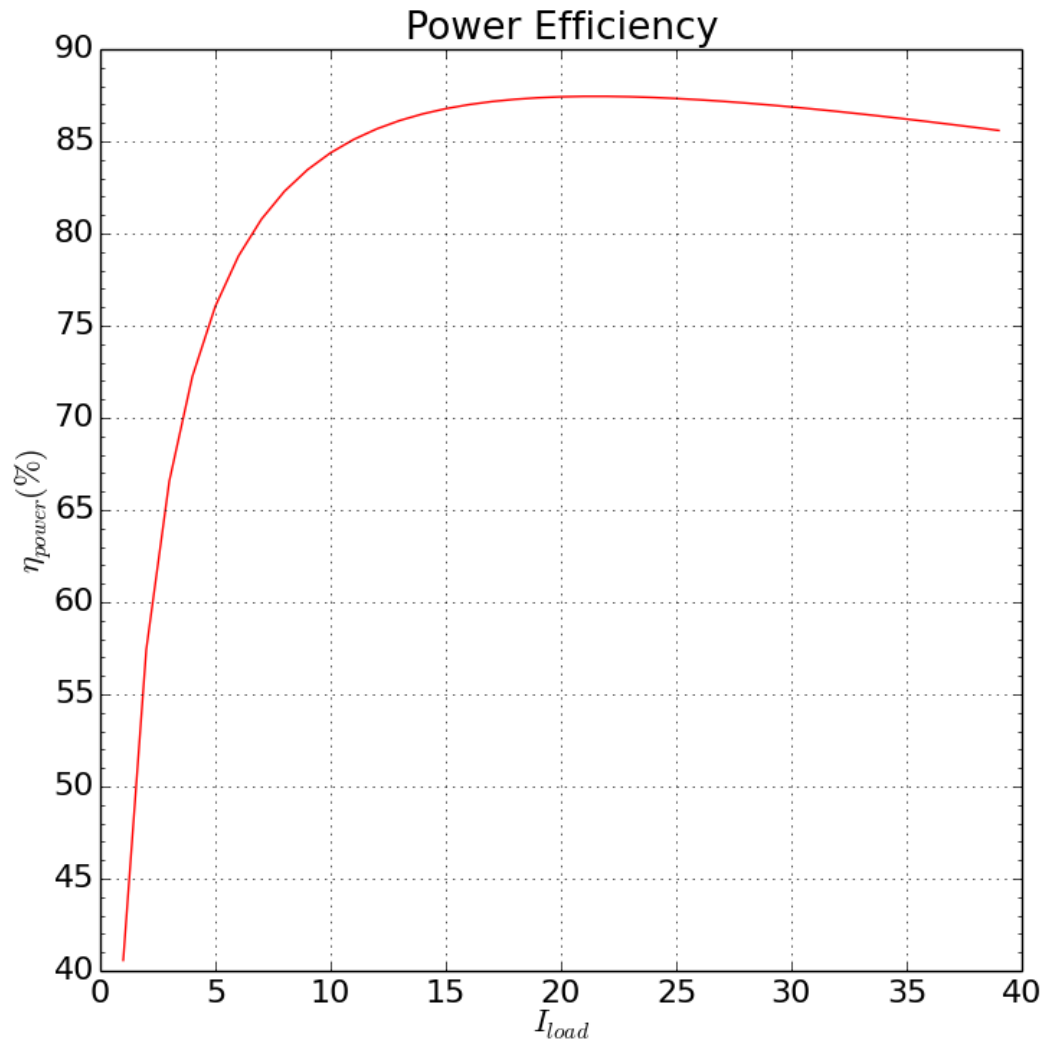


Figure 54: Power efficiency versus output current

As expected high currents cause a decline in efficiency at high currents due to conduction losses. However, a more precipitous decline occurs at low currents. As the current reduces below 11 A, the efficiency drops below 85% and approaches 40% a

decade lower. In fact, the solution covers less than half of a decade as shown in the semi-log plot of Figure 55.

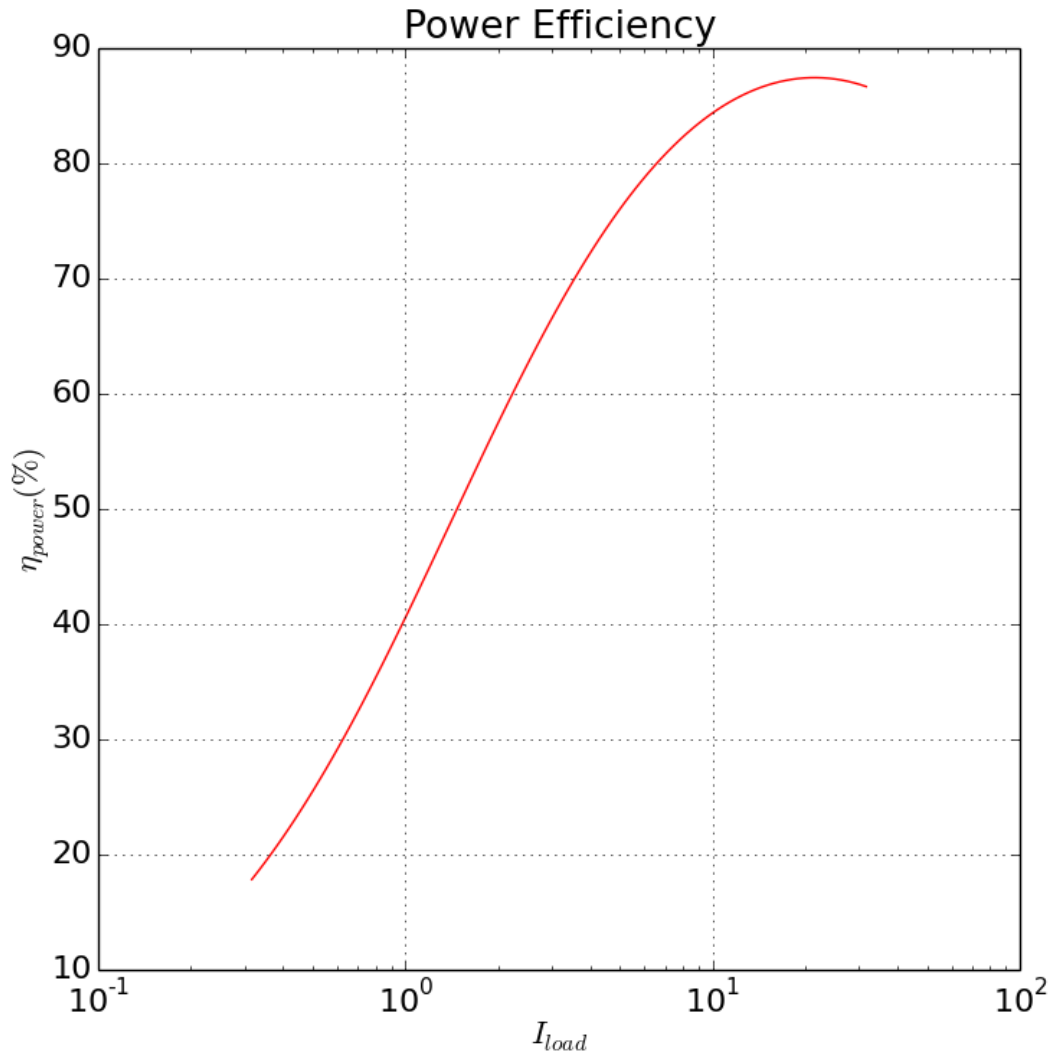


Figure 55: Semi-log plot of power efficiency versus current

The sharp decline in efficiency occurs as a result of the fixed capacitance delivering current through the power FETs. One commonly implemented technique to lift

the efficiency at lower loads is to retire a portion of the power FETs beyond a threshold. At lower currents the conduction requirements of the FETs relax, allowing for a reduced amount of capacitance switched every cycle. However, under this scheme the buck may dive into discontinuous current mode (DCM) or have negative currents running through the inductor. Negative currents should be avoided for long periods, if possible. This detracts from the efficiency since current actually runs negative from the load back to the regulator. This is not catastrophic, but does impact the efficiency. On the other hand, DCM requires a complete control law change, likely pulse frequency modulation (PFM) or similar technique.

The simple solution disables one or more phases as needed as the current drops. This maintains the same control law, ripple, and frequency at the lower currents, minimizing the inherent risk in control law and frequency changes. However, nothing this good comes for free. The n buck-phases must be offset in phase by $2\pi/n$ for proper multi-phase summation and cancellation. If n drops from 8 to 7, the master clock source for the phases must switch from $\pi/4$ separation to $2\pi/7$. Every time the number of phases deviates the PLL or clock divider must make an adjustment to accommodate it. This can be an expensive feature in the clock source, and may incur latency to reacquire lock or rearrange oscillator stages without a power penalty.

In order to avoid reacquiring lock in the clock source, one option is to take advantage of powers of 2. For an 8-phase design with $\pi/4$ separations, elimination of the odd phases takes place without impacting the clock source. In the same vein, a reduction from 4 to 2 phases and 2 to 1 all use the same clocking scheme. Figure 56 shows the power efficiency versus current for this scheme.



Figure 56: Improvement in efficiency through phase shedding

The improvement for phase shedding is abundantly clear in the efficiency plot. With this simple architectural feature the range for which the regulator achieves greater than 85% efficiency increases dramatically. Three steps of shedding expand the usable high efficiency range by a full order of magnitude. The rule for shedding in this case is simple. At one-half of the maximum capacity, phases 1, 3, 5, and 7 shut off. At one-

quarter capacity, phases 2 and 6 disable. Below one-eighth the full output current capacity, a single phase delivers the current to the load.

5.7 Summary of Architecture

The early planning and architectural development of an integrated power supply starts with feasibility study with the inductors, capacitors, and power FETs of Chapter 4 in mind. The load must be bracketed and understood in order to properly specify the voltage regulator. Minimum and maximum currents, ramp rates, and capacitances must be understood. Without fully understanding the load circuit the power supply cannot be specified.

Following the feasibility and load investigation, a normalized phase of the power supply can be specified. This model accepts parasitic data from the Chapter 4 elements and applies basic buck converter equations to architect the power supply.

Voltage droop calculations must be included in the architectural study to guarantee that the regulator does not allow the voltage to exceed the load requirements. If the voltage exceeds the maximum value, transistors in the load may fail prematurely. If the voltage dips below the minimum value, speed paths may fail cause a logic error. AVP aids in meeting these specifications, and must be included in the integrated regulator design.

6 CIRCUIT DESIGN

Basic elements and architectural capabilities open the door for the regulator to move onto the microprocessor package, implying that the controller will reside on the same die to take full advantage of IVR. This chapter discusses the controller and circuitry resident on the microprocessor die.

The input supply has been determined architecturally from chapter 5 to be 2.0 V as a consequence of the maximum transistor voltage and stacking capability. It is nearly impossible to connect the power supply directly to the main input (12-24 V for servers). Package dielectrics begin to break down above 6 V. Stacking extends the usability of native transistor voltages through protection as outlined in Section 4.3.3. Other rails exist on system boards as well, but none of them have sufficient power capacity for a high performance multicore processor. Therefore, the board needs an intermediate stage to regulate sufficiently down so that the high performance transistors can tolerate the input voltage. However, this redistribution stage has reduced regulation requirements since the output is not directly consumed by the microprocessor.

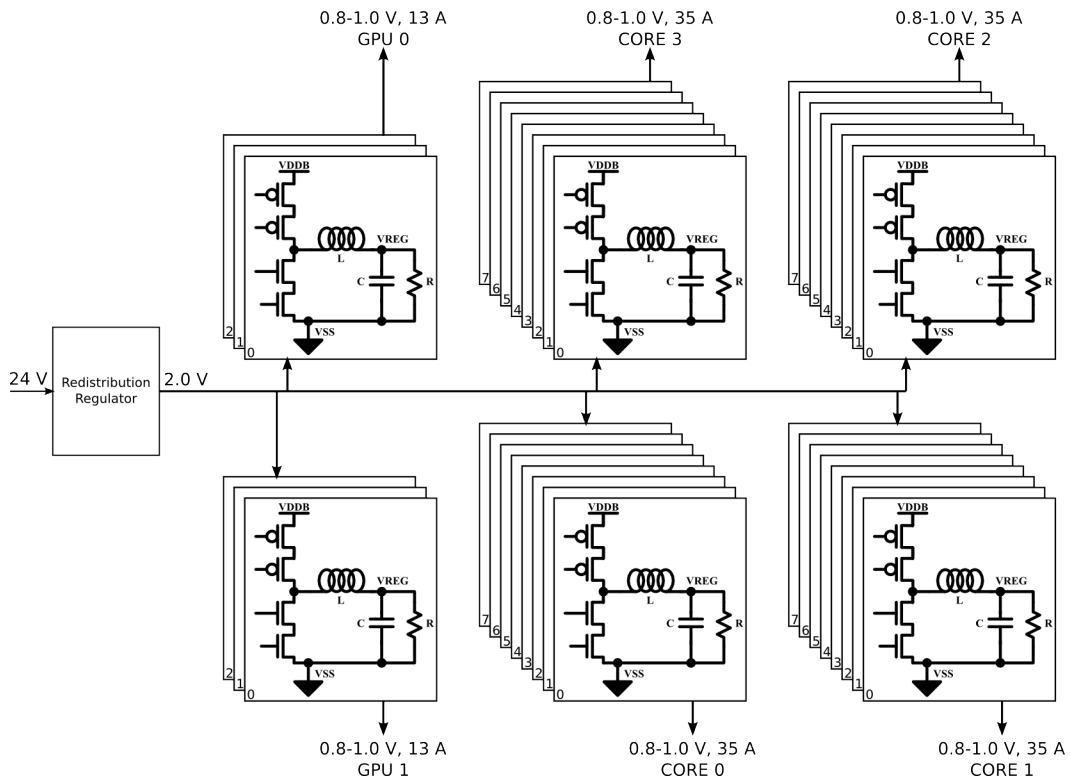


Figure 57: The complete IVR system.

6.1 Controller Strategy

With the phase-based architectural approach, each phase supports a maximum output current. Multiple phases can combine, as needed to drive a single rail. While it is possible to deliver rails with a single phase, limiting the minimum number of phases to two provides lower input and output ripple, better droop performance, and less control circuitry. A controller that can easily gang phases together provides a flexible solution.

The most basic converter controls the output voltage by using an error amplifier, ramp generator, and comparator as shown in Figure 58.

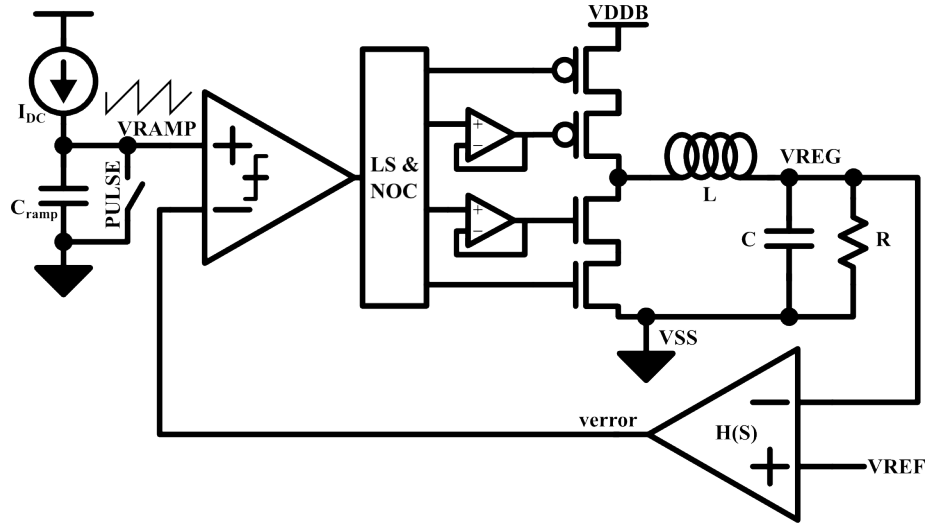


Figure 58: Voltage mode controlled buck topology.

The loop operates by comparing *verror* to the ramping voltage. At the start of the cycle, *vramp* is 0 V, and the comparator output is logic 0. This causes the PMOS power FET to turn on. The inductor begins integrating the voltage across it. Once the ramp crosses *verror*, the comparator drives high, putting the power FETs into synchronous rectifier mode. The *verror* signal is driven up or down by the error amplifier to minimize the difference between *VREF* and *VREG*. The transfer function *H(s)* provides compensation for the loop stability. If *VREG* goes down, *verror* goes up, which causes the comparator to produce a larger duty cycle. The voltage ramp circuit operates via a DC current source feeding a capacitor. The voltage ramps at a rate of I_{DC}/C_{ramp} , and the current required for the ramp to reach the target voltage, V_x , is $V_x * C_{ramp}/T_{per}$.

This controller works well for single-phase solutions. However, each additional phase requires external circuitry to balance the current between phases. Stability, in some cases, can also be challenging. The double pole in the LC filter and the resistive load follows the well-known second order response:

$$H(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \quad (41)$$

The natural frequency, ω_n , is simply $1/\sqrt{LC}$ and the damping ratio, ζ , denotes how steep the phase drops to 180° .

$$\zeta = \frac{1}{2R} \sqrt{\frac{L}{C}} \quad (42)$$

The damping ratio shrinks with reduced inductance. Conventional buck converters enjoy a damping ratio amenable to stabilization using type II or type III compensation because huge capacitors accompany the inductor. Some regulators even rely on the ESR zero to help stabilize the loop. However, because of the physical constraints the damping ratios in IVR can become very small, and the ESR zero is pushed to a very high frequency given the location of the capacitor with respect to the loop, eliminating its use as a stabilizer. A very small damping ratio, ζ , can be difficult to stabilize under all conditions using conventional voltage mode loops.

Current mode control addresses these problems naturally. The current mode topology is depicted in Figure 59.

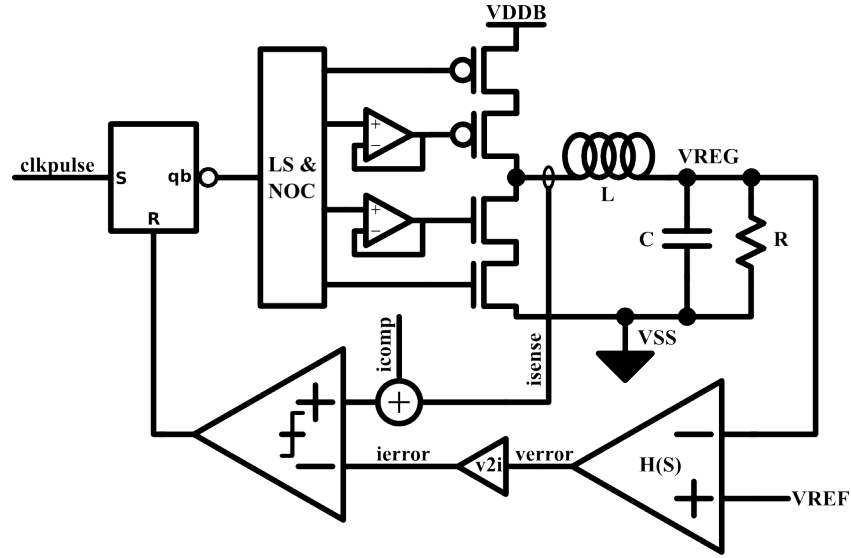


Figure 59: Current mode controlled buck topology.

Current mode control implies that two loops dictate the system output. An inner loop ensures that the peak current through the inductor, i_{sense} , is equal to a target current set by the other loop, i_{target} . The outer loop serves i_{target} until the output voltage, V_{REG} , equals the reference voltage, V_{REF} . Cecil Deisch introduced the topology in 1978 [16]. Compensation is required in the inner loop, labeled i_{comp} , to both prevent sub-harmonic oscillations and outright instability. The problem with the current-loop buck is that solutions to $V_{REG}=V_{REF}$ are not singular. The system can be cyclostationary over one or more clock cycles. For example, if the duty cycle required is 60%, then the buck may oscillate at a frequency of $f_{sw}/2$ between 40% and 80% such that the average observed duty cycle is 60%.

With a current-controlled topology, the current sharing among phases automatically happens due to the nature of the inner current loop. Additionally, since the inner loop forces the target current through the inductor, the inductor part of the LC double pole gets pushed to a high frequency, effectively making the inductor look like a

voltage controlled current source. This reduces the order of the loop by one and helps the damping ratio problem in the LCR circuit.

As an added bonus, the current mode controller provides superior input noise rejection. The input noise rejection stems from the fact that the ramp rate of the inductor current during the charge phase is a function of the input supply. The comparator will happily wait as the inductor integrates $V_{DDB} - V_{REG}$. In other words, the loop takes into consideration that the input supply is variable during the charge phase, which improves supply rejection considerably.

6.1.1 Operational Modes

When the regulator operates in two states, PMOS charge and NMOS synchronous rectification, it is said to operate in Continuous Current Mode (CCM). In CCM the current ramps up to i_2 , down to i_1 , and then repeats. The overall current ripple observed in the inductor would be $\Delta i = i_2 - i_1$. Figure 60 depicts CCM operation over two cycles. The current remains above 0 A at all times.

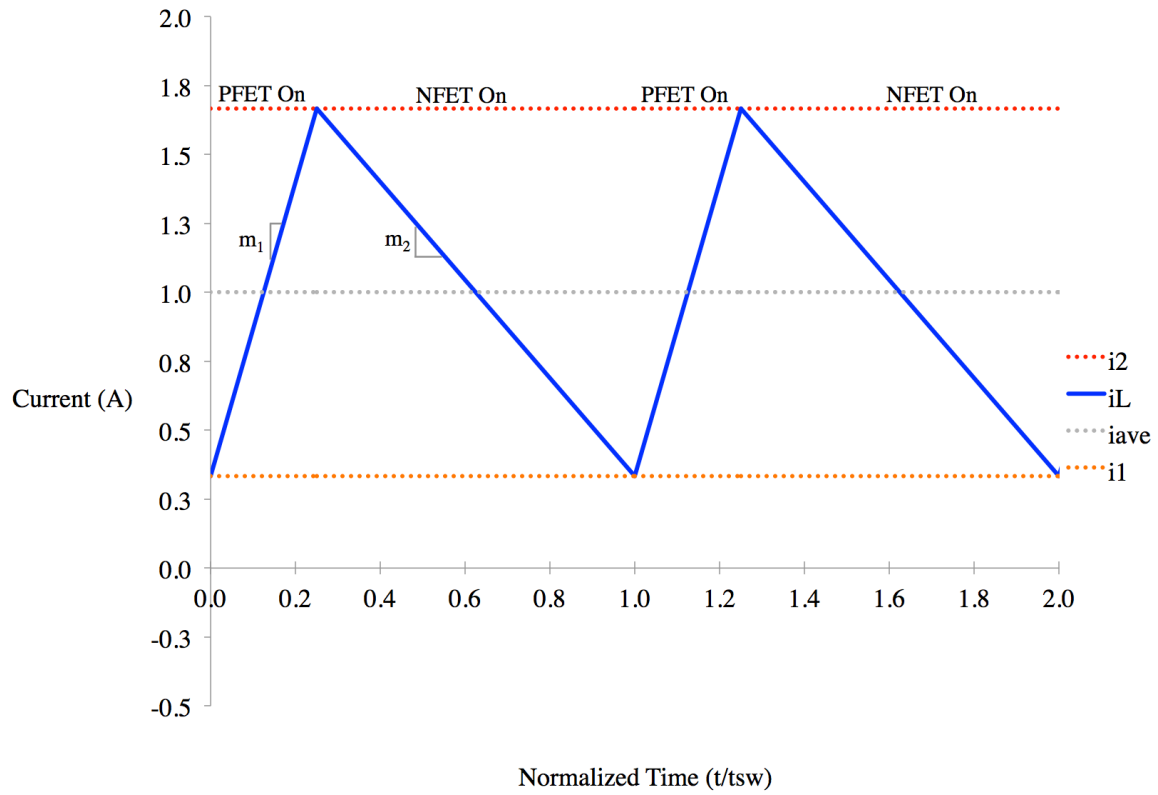


Figure 60: CCM current waveform

If the average current, i_{ave} , drops below $\Delta i/2$, then the current through the inductor becomes negative without a control law change. This may be referred to as negative CCM, shown in Figure 61. Negative CCM is not a desirable operational mode as the unnecessary negative current contributes to the overall loss of the regulator.

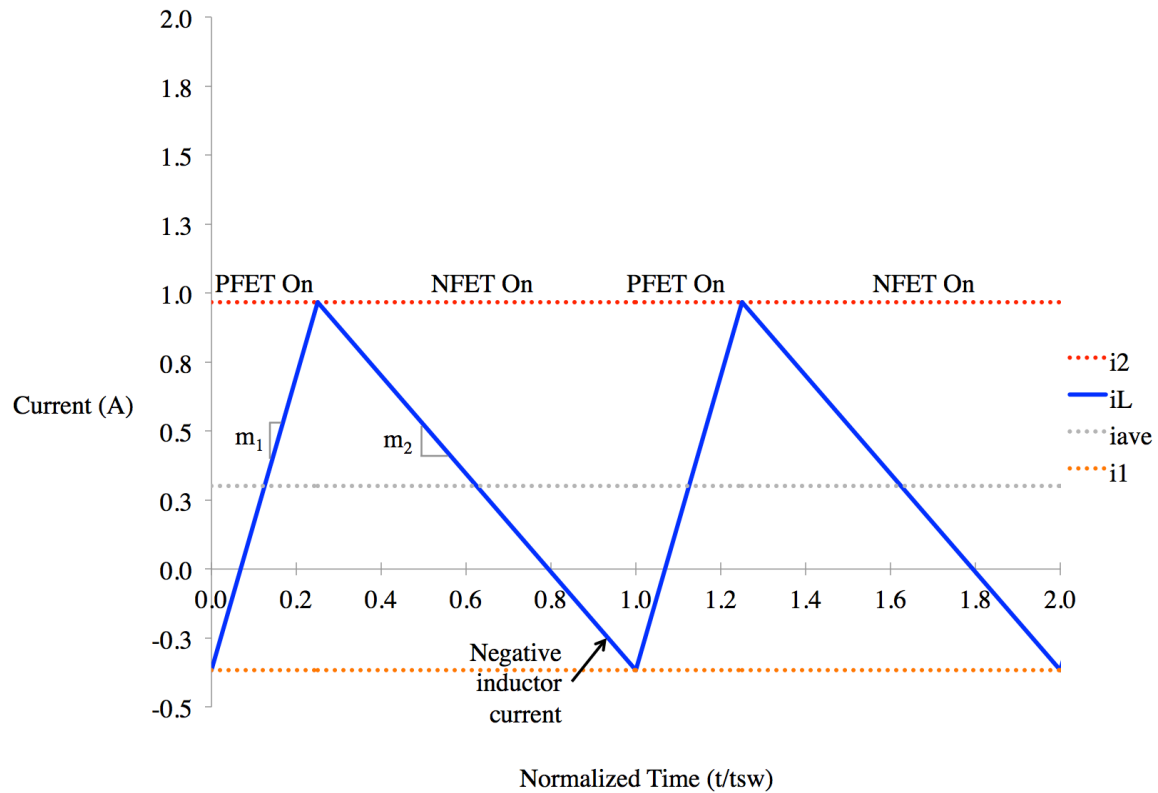


Figure 61: Negative CCM current waveform

The controller can enter tristate instead of allowing the inductor current to become negative. With both NMOS and PMOS power FETs off, no current flows through the inductor. The output load relies on the capacitor alone to maintain the output voltage. This mode is called Discontinuous Current Mode (DCM). Under low loads, DCM can be very efficient since the switching frequency can reduce with the load, lowering dynamic power along with the load current. However, changing modes requires overarching control to determine which mode the circuits operate in, and when to change over. Figure 62 illustrates DCM current waveforms across two cycles.

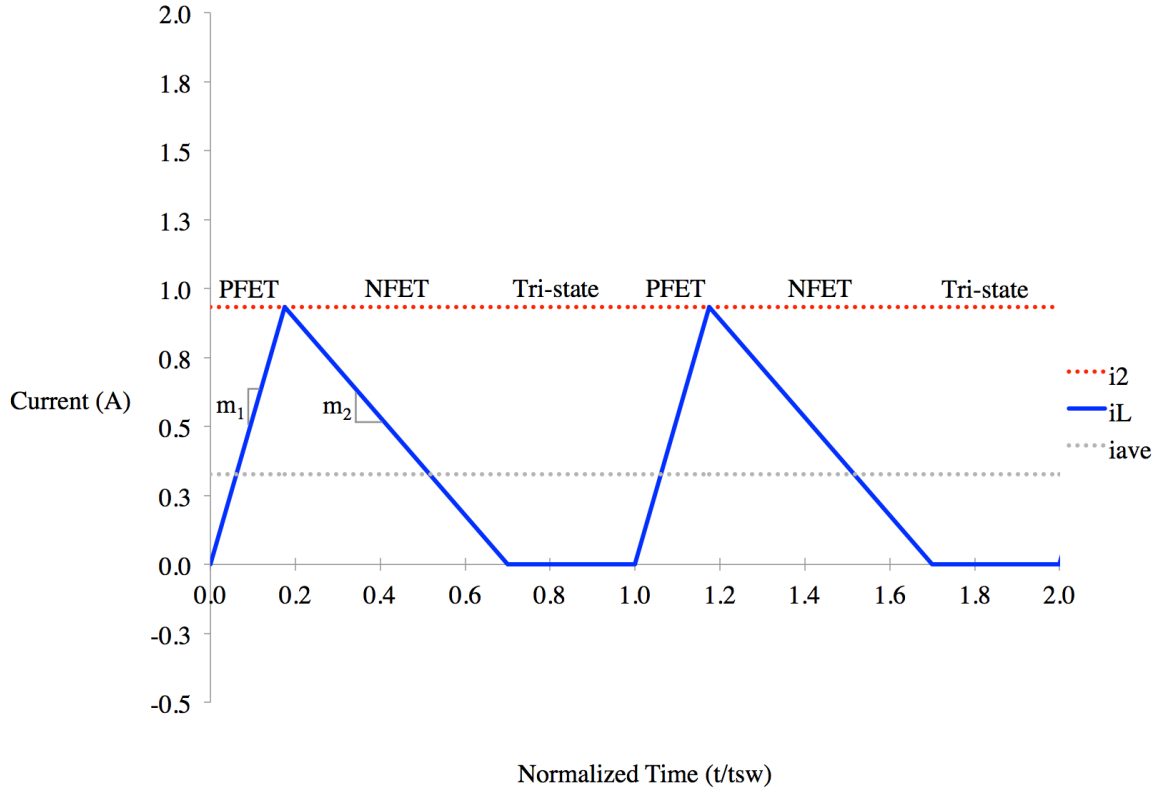


Figure 62: DCM current waveform

In DCM T_P , T_N , and T_Z define the full cycle. T_P the duration while the PFET is on. T_N represents the duration for which the NFET is on. Finally, T_Z represents the high impedance duration. The sum of T_P , T_N , and T_Z must equate to T_{switch} , the period. The average current can be determined from simple geometry.

$$I_{\text{average}} = i_2 \frac{T_P + T_N}{2 \cdot T_{\text{switch}}} \quad (43)$$

Control law changes must be guard-banded with hysteresis. Hysteresis exploits the fact that the converter will typically operate in a narrow band for most cases. If the operating region is at the boundary of the law change, then hysteresis prevents the controller from constantly changing states, requiring a larger change in the operation

before changing states. However, if the operating region is as wide as the hysteresis band then the system can become unstable. This may be referred to as “motor-boating.” AC loop stability does not reveal large-scale instabilities. Stability analysis typically involves a small AC perturbation around a DC operating point, assuming that the DC operating point remains constant. However, if the operating point itself changes sufficiently, then the small AC perturbation does not adequately describe the loop.

Only a transient simulation can reveal large signal stability problems. Consider an amplifier loop in an AC simulation. The simulator finds the DC solution at a given region of operation for all devices. If the devices can enter more than one region of operation then a large signal transient simulation must be used to determine whether or not the circuit is ultimately stable.

Luckily, for Tuolumne and microprocessors in high performance technologies, the minimum currents are still quite high due to leakage and sub-threshold currents. Discontinuous current mode and additional control laws are avoided altogether. For small supplies this would need to be reconsidered, possibly adding in a second control law.

6.2 Block Circuits

With the regulator controller, architecture, and topology in place, block level circuit design can be completed.

6.2.1 Power FETs

The power FETs deliver current to the inductor from the input supply during the charge phase, and maintain a low impedance path for the current during synchronous

rectification to drain energy stored in the inductor's magnetic core. The power FET array dominates the integrated circuit area of the converter.

As outlined in section 4.3, the V_{DDB} input supply voltage extends beyond what a single transistor can tolerate, requiring a cascode structure to keep V_{DS} and V_{GS} within the safe operating range of the PFET and NFET.

Figure 63 depicts the power FETs within the dual loop controller. The switching gate signals arrive from the level shift and non-overlapping clock (LSNOC) generator block, which splits the pulse-width modulation signal into the proper gate voltages.

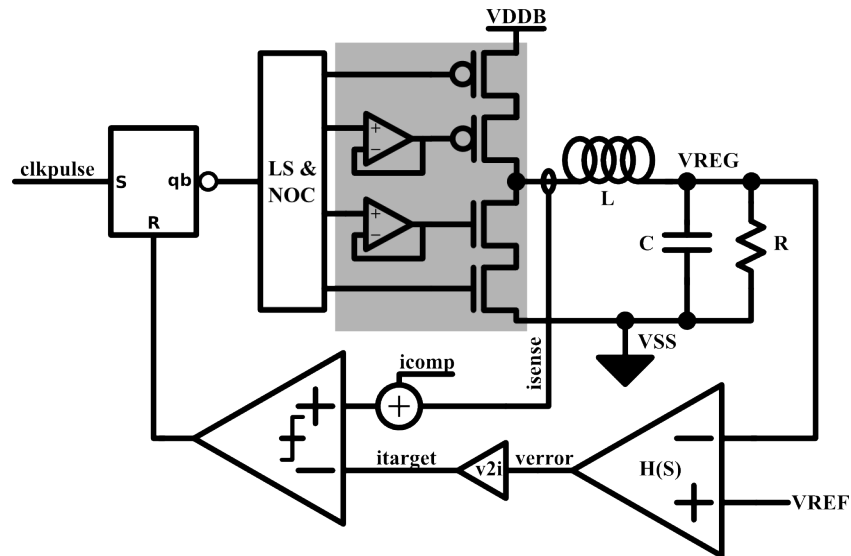


Figure 63: Power FETs within current mode loop.

Two amplifiers bias the gates of the inner FETs according to fixed reference voltages. These amplifiers keep the impedance at the gate of the cascode transistors, MCP and MCN, very low at DC. The capacitor, C_0 , takes over this job at high frequencies. The cascode amplifier structures are shown in Figure 64.

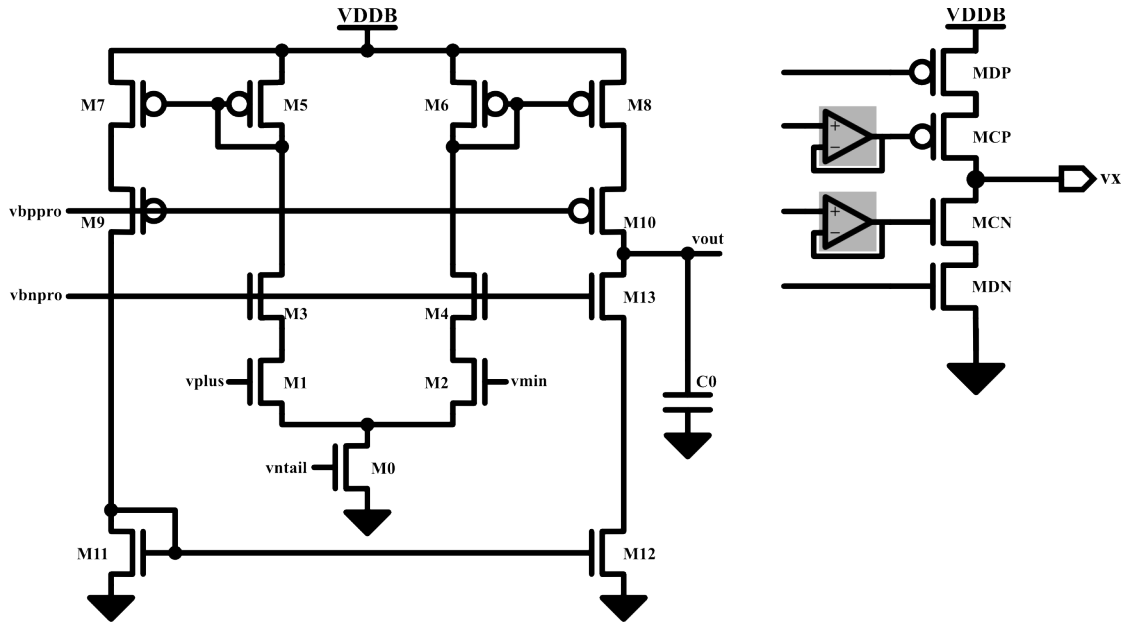


Figure 64: Schematic of power FET MCP and MCN gate-bias amplifier

6.2.2 Level Shifter and Non-overlapping Clock Generator (LSNOC)

The LSNOC performs several digital functions in one block. This block receives a single PWM signal from the loop, and buffers the gates of the final power stage of the converter. Figure 65 illustrates the LSNOC within the overall loop.

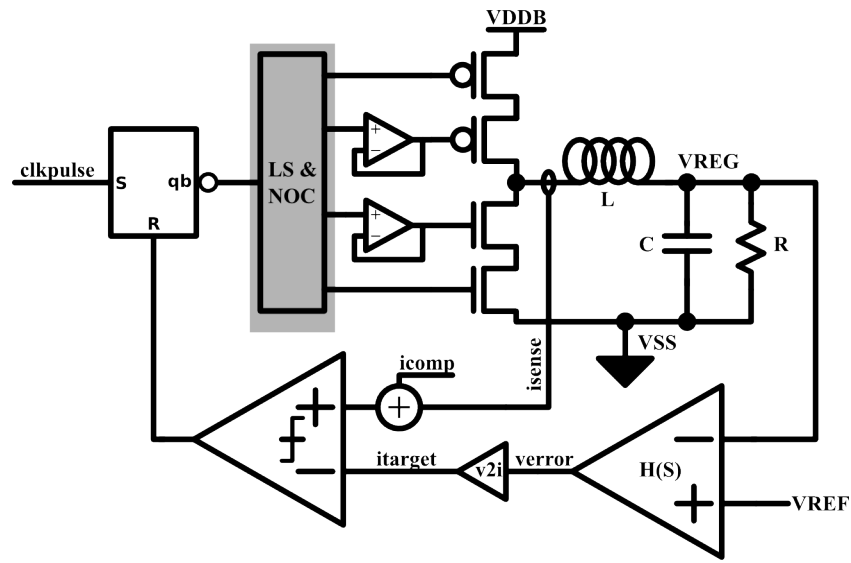


Figure 65: LSNOG within current mode loop

The most important job of the LSNOG is the non-overlap function, also known as break-before-make. A simplified logic diagram of the non-overlap circuit is depicted in Figure 66.

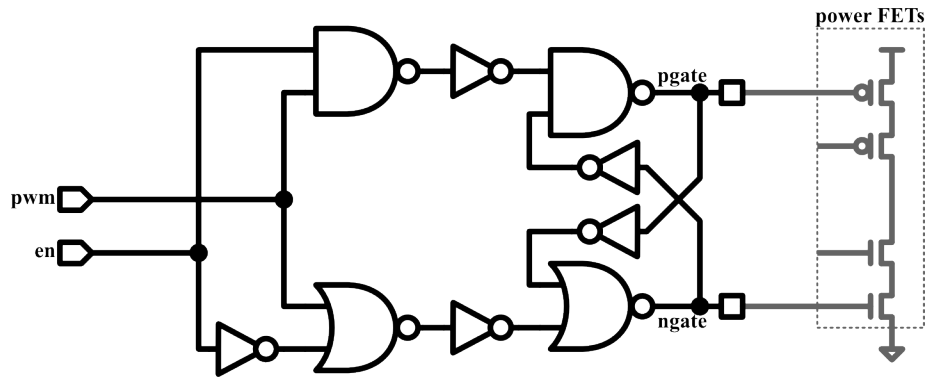


Figure 66: Non-overlap logic

The first set of NAND and NOR gates allows the output stage to be in tri-state, disabling both PMOS and NMOS power FETs. The cross-coupled stage of NAND and NOR gates comprise an SR latch, which performs the non-overlap functionality. The

ngate signal cannot be driven high until the pgate signal goes high. Similarly the pgate signal cannot be driven low until the ngate signal goes low. The break operation occurs uninhibited. Delay in the feedback inverters allows for tuning of the delay.

The full circuit for the non-overlap circuitry is shown in Figure 67.

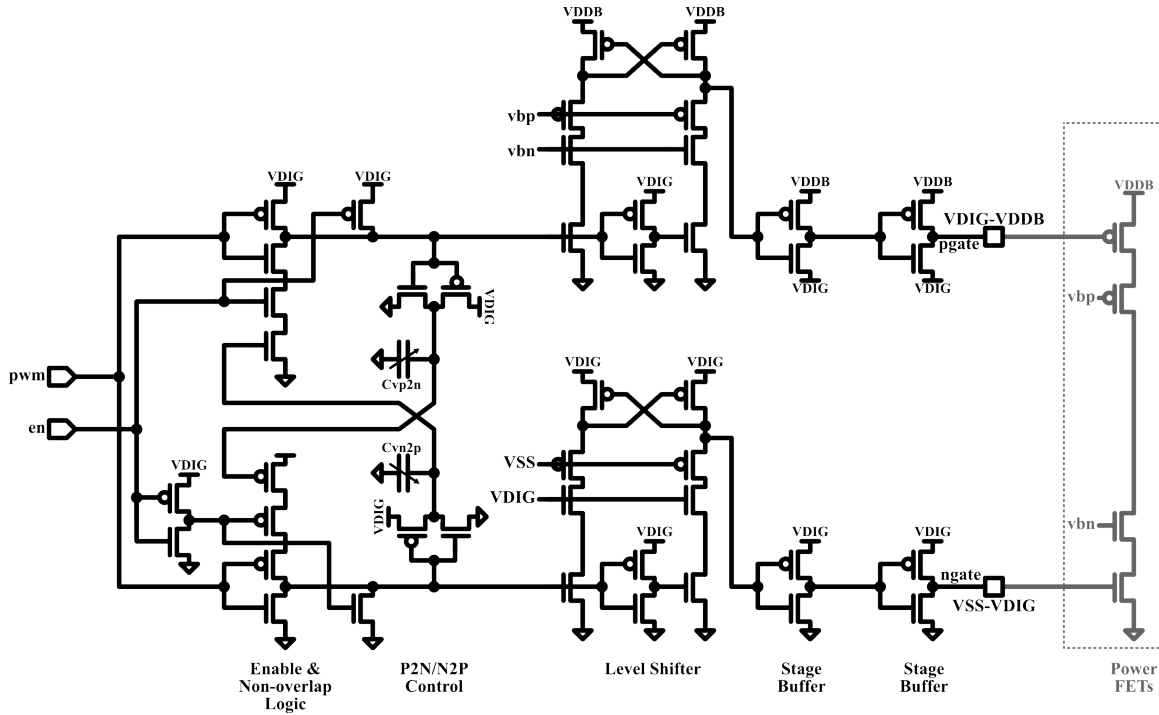


Figure 67: Level shift and non-overlapping tapered buffer stage circuit

Two stages of buffers provide gain for the large capacitive load on the gate of the power FETs. As described in Chapter 4, the PMOS gate buffers switch between VDDDB and VDIG, and the NMOS gate buffers toggle between VDIG and VSS. One level shifter is required to level shift from VDIG to VDDDB-VDIG. However, a second dummy level shifter provides skew relief in the timing for both PMOS and NMOS paths where the NMOS path doesn't need a level shifter. The enable and non-overlap logic are combined into a single stage on the front end. Capacitors Cvp2n and Cvn2p are digitally

programmable. The load the feedback inverters, modulating the crossover delay for PMOS-to-NMOS and NMOS-to-PMOS independently. Without sufficient separation, large crowbar current may develop directly between V_{DDB} and V_{SS} , wasting tremendous energy. Too much delay causes power to be wasted in the NMOS rectifier in the off state while the inductor forces current through it.

6.2.3 Current Sensor

The current sensor performs one of the most important functions within the current mode loop, and does not exist in voltage mode loops.

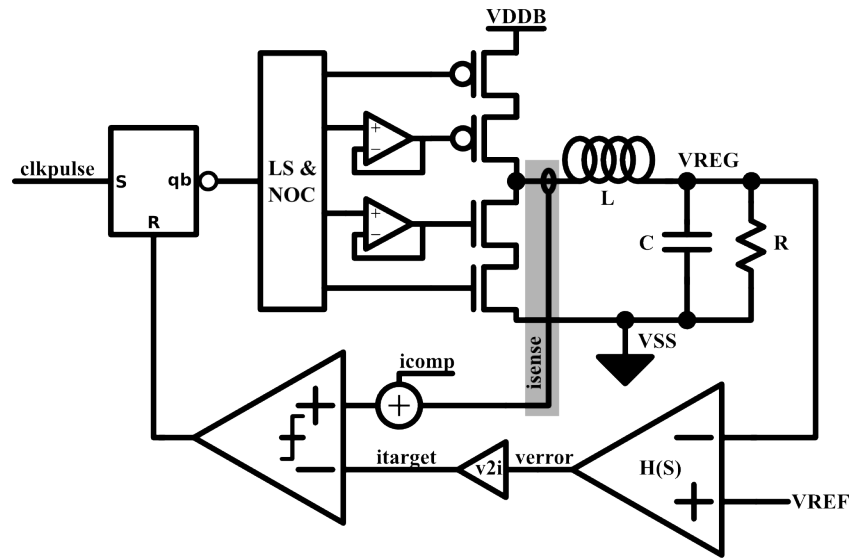


Figure 68: Current sensor within current mode loop

Many different techniques have been applied to provide a proportional signal that represents the inductor current. A series milliohm resistor provides the most accurate signal, but this is not a viable solution in IVR scenarios. Another solution involves exploiting the DC resistance of the inductor by filtering the voltage across it, thereby

producing a voltage proportional to the current. However, this method is prone to temperature instabilities, as the DC resistance of the inductor possesses a large temperature coefficient, corrupting the signal. Yet another tactic involves measuring the voltage across the PMOS power FET directly. Similar to the temperature problem in the inductor, VDS of the PMOS device varies substantially.

One other alternative is to implement a dummy PMOS sensor, identical to the power stage. An amplifier servos the current through the dummy stage until the voltage across it and the main power stage are identical. Shown in Figure 69 is a topology that suits IVR.

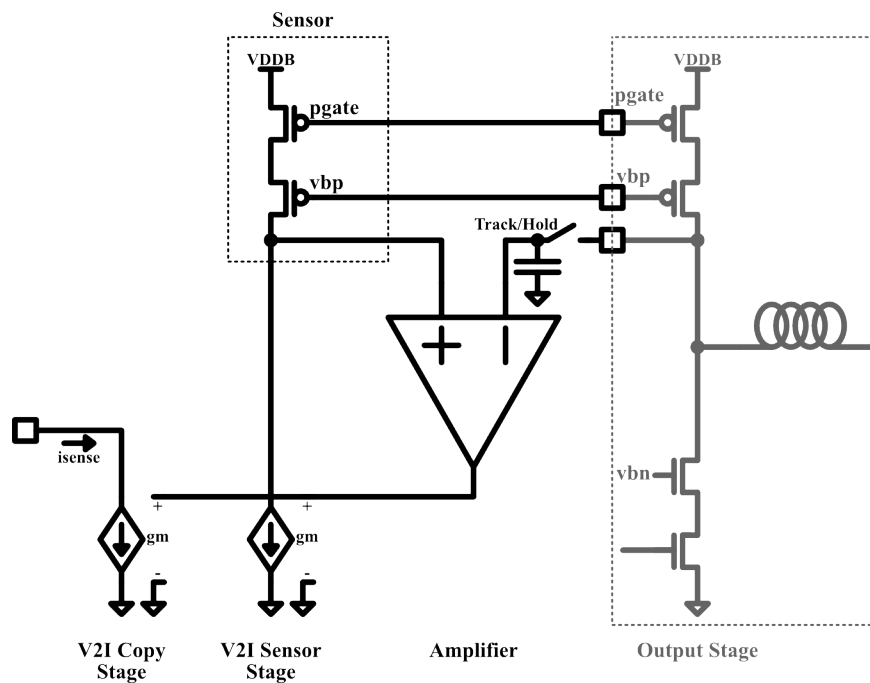


Figure 69: Current sensor topology

Several issues require careful circuit design to achieve an accurate representation of the inductor current. Given that the voltage across the power stage is made very small

for efficiency purposes, the input common mode of the amplifier is essentially a few millivolts from the supply level. Note that a track and hold circuit must hold the approximate state of the input to the inductor during synchronous rectification. Otherwise the amplifier input common mode would drive towards ground, causing the amplifier to go out of saturation and incur a latency to reacquire the signal.

Given that the power stage and sensor are by definition very low impedance, gain of the V2I sensor stage can be challenging. In fact, the sensor stage works best as a source follower that tries to keep from losing all the gain rather than adding to it. The schematic of the current sensor is shown in Figure 70.

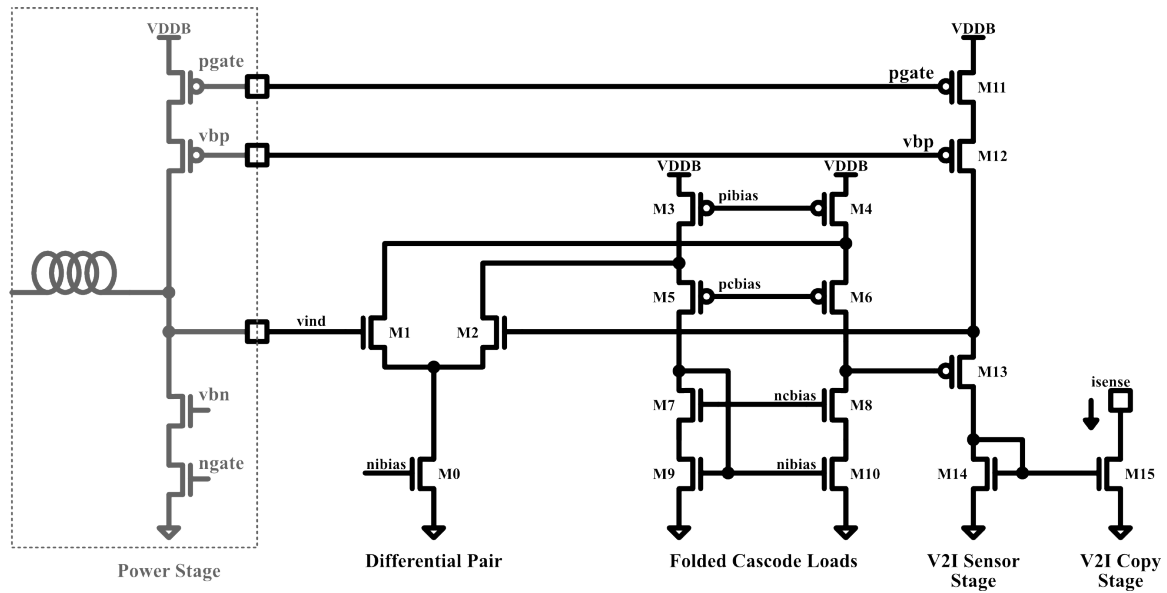


Figure 70: Schematic of current sensor circuit (track and hold not shown)

Note that the gates of M11 and M12 are identical to the power stage for best matching. However, when the converter switches over to the NMOS side the pgate signal goes high, turning off the V2I sensor stage. This nulls the output of the sensor stage. Once the converter is in the NMOS cycle, the output of the current sensor will not be

6.2.4 Current Loop Stabilizer

Creating a ramping current waveform proves to be challenging when compared to a voltage ramp waveform. A DC current driven onto a capacitor with a periodic shunt

reset provides a perfect voltage ramp waveform. A DC voltage driven into an inductor with a periodic shunt would provide the perfect current ramp waveform. However, inductors on the IC are out of the question due to their large size. So, a simple voltage ramp waveform is created. Then a V2I amplifier loop converts the ramping voltage into a current directly. This circuit is shown in Figure 72.

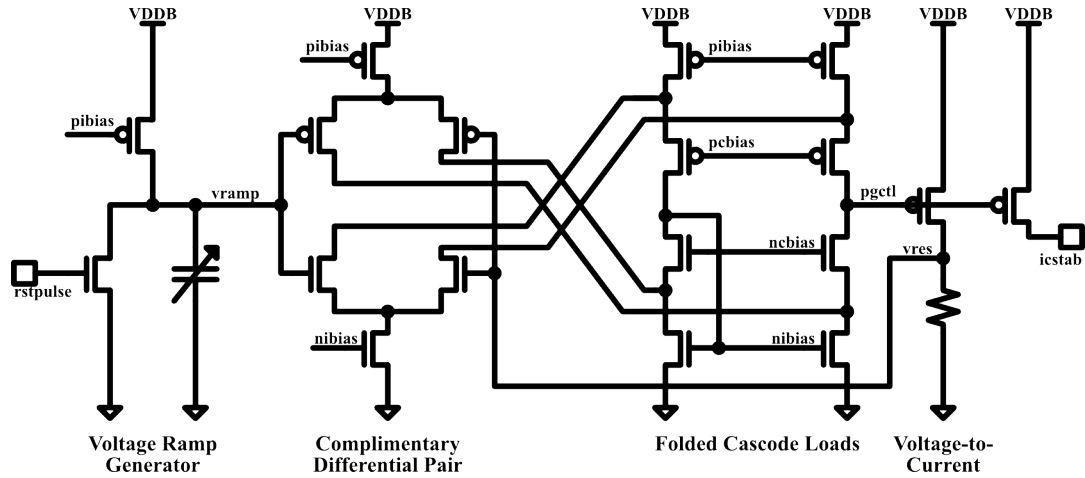


Figure 72: Schematic of current stabilizer

A programmable capacitor allows the ramp rate to be digitally programmable. The complimentary input differential pair allows the ramp to occur over the full VDDB range. Finally, a PMOS common source amplifier drives a current into a resistor to perform the conversion.

6.2.5 Current Comparison Circuit

Second to the current sensor, the current comparator forms the heart of the current mode controller. The comparator combines the outer voltage loop and inner current loop, stabilized, to determine when to switch from PMOS to NMOS in the power FET stage.

The inputs to the comparator are currents, given the fundamental operation of the topology. Other reasons to use a current signal, rather than voltage, appear later. Figure 74 depicts the current comparator within the current mode loop. The current comparator drives the reset pin of an SR latch, which is set dominant.

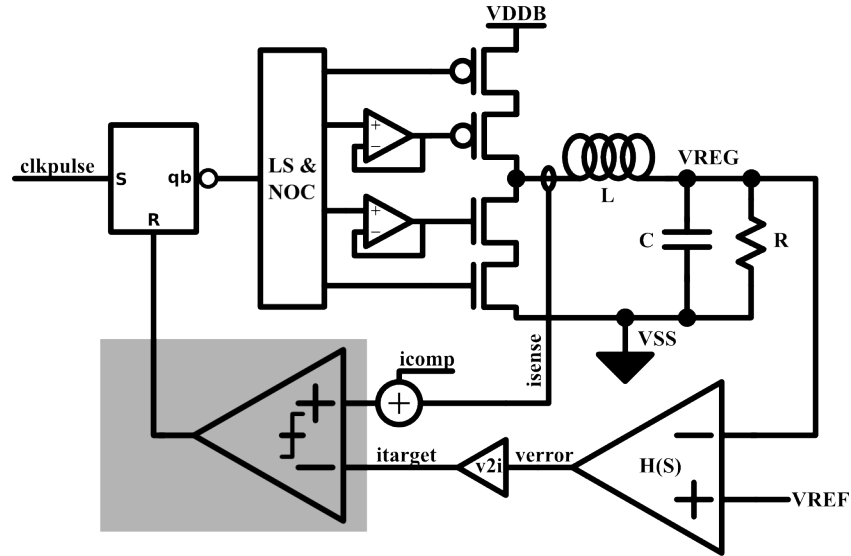


Figure 73: Current loop comparator within current mode loop

The full circuit for the current comparator is shown in Figure 74.

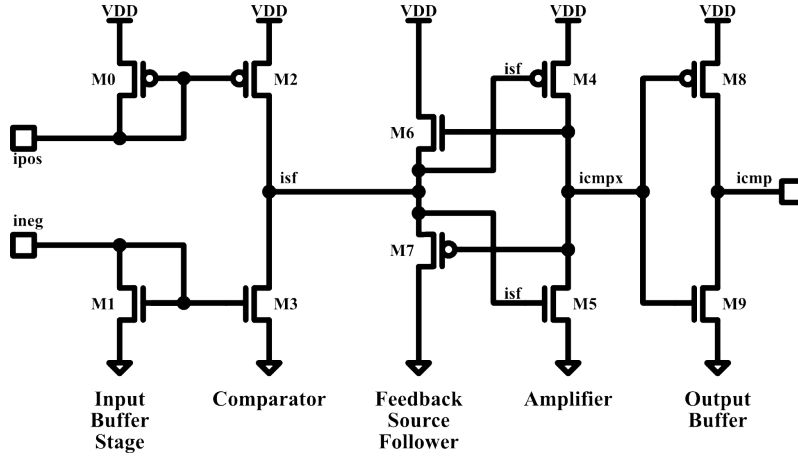


Figure 74: Schematic of current comparator circuit

The ipos input port receives current signals from the current stabilizer combined with the current sensor output. These two signals add together simply by combining the two current sources, an advantage of using current signals over voltage. The ineg input port receives the target current waveform from the voltage error amplifier. These two current ports feed into a mirror stage. The outputs of the mirrors drive each other at a high impedance node, isf. Without any further circuitry, the current comparison occurs at this point. Transistors M4, M5, M8, and M9 simply buffer the signal, providing additional gain.

The isf node, being high impedance and driving a gate capacitance, has a tremendous time constant. This time constant is:

$$C_{inx} \cdot (g_{dsp_M2} || g_{dsn_M3})^{-1} = C_{inx} \cdot r_{out} \quad (44)$$

In order to maintain accuracy, r_{out} may reach as high as 10 M Ω . For a load capacitance of 20 fF, the time constant on this node would be 200 ns. In this design the clock period is 33 ns, rendering the comparator useless. Fortunately, a solution exists. This solution is to use a source follower, in feedback, to lower the impedance of the

comparison node without sacrificing the performance of the current sources. H. Traff published this technique in 1992 in the context of analog to digital converters [59]. This circuit appears often in optical communications links, where the incoming signal is a current from a diode that is to be converted into a digital voltage. It is therefore referred to as a transimpedance amplifier, or TIA.

6.2.6 Voltage Error Amplifier

The final important block within the current mode topology is the voltage error amplification stage. This stage reads the output voltage, compares it to the reference, and then delivers an output current that is the target for the current comparator. If the error signal is negative, VREG being greater than VREF, then the target current decreases. If the error signal is positive, VREG being less than VREF, then the target current increases. This outer voltage loop is shown within the current controlled loop in Figure 75.

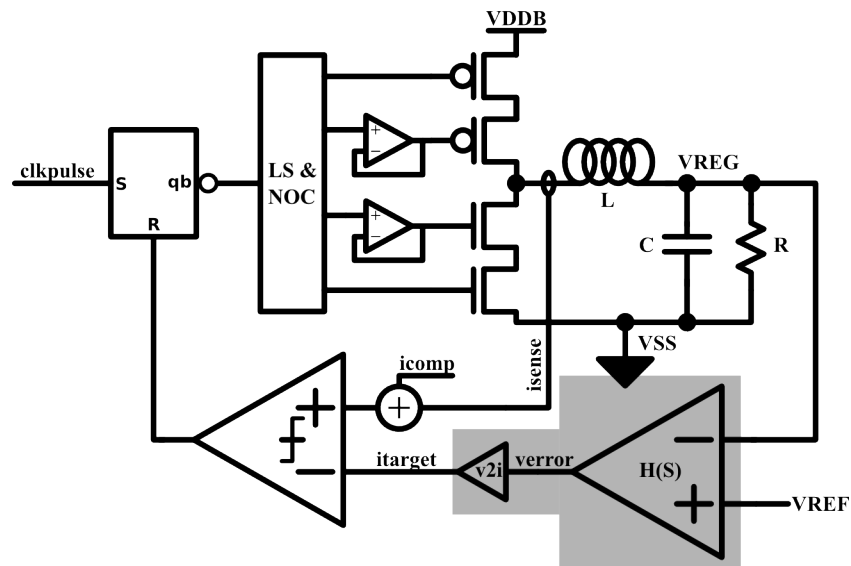


Figure 75: Voltage error amplifier within current mode loop

power supply. Left unchecked, this noise can damage the transistors, cause stability issues, or even trickle through the regulator to reach the sensitive digital circuits. However, this job becomes simpler for the system designer since IVR reduces the number of supplies coming into the chip into just one.

The second reason for input supply filtering is for electromagnetic interference (EMI) reasons. The cyclical behavior of the buck on the input supply may cause a disturbance in other parts of the system.

6.4 Circuit Summary

The circuits presented throughout this chapter together form an IVR solution. These circuits are all CMOS compatible, with no external elements required. The following circuits comprise the current controlled integrated regulator:

- Power FETs
- Level shifter and non-overlapping clock generator
- Current sensor
- Current loop stabilizer
- Current comparison circuit
- Voltage error amplifier

In addition to these integrated circuits and the fundamental elements of Chapter 4, input supply capacitance provides decoupling of the high frequency noise generated on the input supply when the current switches on and off.

7 CONCLUSIONS

Table 9 summarizes the performance of the conventional and integrated regulator solutions. The conventional design data comes from the Intel voltage regulator design guide [28] referred to in section 2.6. The integrated solution implements the elements, architecture, and circuits of chapters 4-6 in an implementation that meets the Intel voltage regulator guidelines in [28].

Specification	Symbol	Conventional	Integrated
DC input voltage	$V_{DDDB_{DC}}$	12 V	2 V
DC output voltage	$V_{REG_{DC}}$	1.1 V	1.1 V
Nominal duty cycle	D	9.2%	55%
Total multi-phase current	I_{total}	180 A	180 A
Per phase capacitance	C_{ϕ}	200 μ F	4.4 μ F
Power efficiency	η_{pwr}	70-80%	85%
Load line resistance	R_{ll}	0.8 m Ω	0.2 m Ω
Output voltage tolerance	$V_{REG_{TOL}}$	220 mV	50 mV
Voltage ripple	V_{ripple}	10 mV	1 mV
Number of phases	n	4	42
Max phase current capacity	$I_{avephase}$	45 A	4.3 A
Source area	A_{src}	4418 mm ²	392 mm ²
Load area	A_{ld}	450 mm ²	450 mm ²
Area ratio	A_{src}/A_{ld}	10.0x	0.9x
Switching frequency	f_{sw}	750 kHz	30 MHz
Inductance	L	470 nH	27 nH
Phase delta-I	Δi	2.2 A	0.6 A
Charge current slope	m_1	23.2 MA/s	33.0 MA/s
Discharge current slope	m_2	-2.3 MA/s	-40.3 MA/s
Source power density	SPD	0.045 W/mm ²	0.459 W/mm ²
Source current density	SCD	0.041 A/mm ²	0.505 A/mm ²
Feedback latency	τ_{fb}	1-2 μ s	< 20 ns

Table 9: Comparison of conventional and integrated voltage regulators.

As the head to head comparison of Table 9 clearly illustrates, the integrated regulator can accommodate or exceed the Intel design guidelines. The conventional

regulator is penalized for being farther away from the load. Advantages and disadvantages both exist for migrating to an IVR solution.

7.1 Advantages

Total capacitance is reduced for the IVR case, which saves on bill of material (BOM) costs as well as area. Additionally, the ratio of source and load areas shows a significant improvement for IVR. The IVR regulator is very similar in size to the load that it supplies whereas the conventional solution occupies ten times the area of its load.

The power advantage of IVR is realized in many different ways. Raw power efficiency of IVR wins due to the superior switching performance of the on-die transistors, for example. However, the voltage tolerance is also $\sim 4\times$ tighter than the conventional guideline. The voltage droop analysis lines up with this perfectly as the IVR load line would produce a voltage droop close to the theoretical 50 mV point from section 5.4. This suggests that the supply could actually be reduced at least 150 mV to reduce power. However, the efficiency alone does not reflect these savings.

The IVR solution provides 42 unique phases with which the architect may choose to slice into 42 individual power islands. Operating with a single phase does not take advantage of multiphase cancellation, though. An optimal solution would slice the rail into 21 unique phases of 2 inductors apiece and the appropriate portion of capacitance per phase.

Another important advantage for IVR is in the feedback latency. The on die analog circuitry which is directly adjacent to the load is limited only by analog design capabilities rather than a long winding feedback path from a sense location back to the PWM controller on the motherboard. This is a 50x improvement in latency.

One less obvious benefit of IVR is that extra voltage margin is removed from the final product. Die are binned and tested with high accuracy power supplies on Automated Tester Electronics (ATE) during manufacturing. The tested product is then inserted into the final motherboard with its own unique power supply. The tolerances of motherboard power supply and ATE must be added into the final system test. Therefore, additional voltage margin has to be added on during test to make sure that the part will work on a supply that could be lower than the high accuracy tester setup.

Once IVR is in place, the part is tested with the final power supply that will power the chip in the end product, which allows the voltage to be aggressively lowered. The lower voltage can add frequency to the part or reduce power.

7.2 Disadvantages

A drawback to the proposed IVR solution is the 2 V input. Modern server racks and infrastructure support a 12 V output from the rectifier that would see a 6x increase in output current with the 2 V rail. Another potential hazard includes the input supply to the buck regulator. Modern package resonances see an increase in impedance around 10-500 MHz range, and the proposed IVR solution would perturb the input supply in that frequency range. A careful analysis of the signal integrity of the input supply must occur to properly supply the IVR solution.

As with any other critical piece of IP brought into the complex SOC world, there are numerous risks and downfalls associated with IVR. These risks include the ability to validate an IC if the power supply fails. There are also unforgiving electrical hazards associated with switching large amounts of current on die rather than in an external device that could be replaced on the board if necessary.

Overall, the benefits of IVR clearly overcome the drawbacks. IVR provides a long-term path for increasing integration. As systems incorporate more IP onto a single piece of silicon, IVR can provide individual supplies to each piece of IP whereas conventional voltage regulators cannot keep up with the integration pace.

7.3 Future Work

Many different aspects of integrated voltage regulation present opportunities for future work. These other areas of future work are explored below.

7.3.1 Fully Integrated Possibilities

Fully Integrated Voltage Regulator (FIVR) implies that the entire regulator is within the microprocessor die including inductor and capacitor. While an architecture using package passives presents a valuable solution to high levels of integration today, there is no sign that integration will slow down. At some point integration and miniaturization forces the regulator onto the die with all the other circuits. There is a lot of work to be done for this to occur in the fields of material science, circuit techniques, and system architecture. The availability of miniature passives opened the opportunity to pull inductors onto the package in this dissertation. Future research in on-die passives will do the same on a much smaller scale.

7.3.2 CMOS and Supply Voltage Departure

As mentioned, the rapid pace of transistor technology leaves behind support for higher voltages that are key for power conversion. For example, the presented solution only supports a 2 V input where the prior work supported 12 V.

In a few cases process engineers have found ways to support higher voltages by drain extension or specially engineered gate oxides. However, there exists a competitive advantage in engineering a transistor specifically for power conversion.

7.3.3 Architectural Features for Low Power

One of the most challenging roadblocks to designing a higher performance power supply is the latency and accuracy of information from the CPU itself regarding its true power state. For example, as the CPU wakes up and begins to service a long piece of code there must be a way to predict an increase in current consumption and relay this information to the regulator for tighter control of the output.

This proves to be a challenging proposition ripe for architectural research. Research work should consider a CPU as a digital to analog converter. The analog information comes in the form of current or power as a result of internal states. The processor needs a method to collect that information and send a code to the regulator representing the current minimizing the latency. The specifications for such a feature would come in the form of any other digital to analog converter: sample rate, bandwidth, and effective number of bits.

BIBLIOGRAPHY

- [1] Allen, Phillip, and Douglas Holberg. *Analog Circuit Design*. 2nd. Oxford University Press, 2002.
- [2] Alon, Elad, and Mark Horowitz. "Integrated Regulation for Energy-Efficient Digital Circuits." *IEEE Journal of Solid State Circuits*, vol. 43, pp. 1795-1807, 2008.
- [3] Ang, Simon, and Alejandro Oliva. *Power-Switching Converters*. 3rd. CRC Press, 2010.
- [4] Arizona State University Predictive Technology Model (PTM). <http://ptm.asu.edu>. September 2008.
- [5] Auth, Chris, et al. "A 22 nm High Performance and Low-Power CMOS Technology Featuring Fully-Depleted Tri-Gate Transistors, Self-Aligned Contacts, and High Density MIM Capacitors." *2012 Symposium on VLSI Technology Digest of Technical Papers*. 2012, pp. 131-132.
- [6] Banba, H., et al. "A CMOS Bandgap Reference Circuit with Sub-1-V Operation." *IEEE Journal of Solid State Circuits*, vol. 34, pp 670-674, 1999.
- [7] Breussegem, Tom Van, and Michiel Steyaert. "An 82% Efficiency 0.5% Ripple 16-phase Fully Integrated Capacitive Voltage Doubler." *IEEE Proceedings of the Symposium on VLSI Circuits*. 2009, pp. 198-199.
- [8] Buller, J. F., et al. "Bandgap Circuit Design Challenges in High Performance 32nm Technology." *IEEE Proceedings of Custom Integrated Circuits Conference*. 2011.
- [9] Chang, Leland, R. K. Montoye, B. L. Ji, A. J. Weger, K. G. Stawiasz, and R. H. Dennard. "A Fully Integrated Switched-Capacitor 2:1 Voltage Converter with

- Regulation Capability and 90% Efficiency at 2.3 A/mm².” *IEEE Proceedings of the Symposium on VLSI Circuits*. 2010, pp. 55-56.
- [10] Chowdhury, I., and Dongsheng Ma. “Design of Reconfigurable and Robust Integrated SC Power Converter for Self-Powered Energy-Efficient Devices.” *IEEE Transactions on Industrial Electronics*, vol. 56, pp. 4018-4028, October 2009.
- [11] Cockcroft, J. D., and E. T. Walton. “Experiments with High Velocity Positive Ions.” *Proceedings of the Royal Society of London*, pp. 619-630, June 1932.
- [12] Coilcraft Incorporated. “Square Air Core Inductors.” 0908SQ Product Datasheet. <http://www.coilcraft.com>. April 2012.
- [13] Dancy, A. P., R. Amirtharajah, and A. P. Chandrakasan. “High-Efficiency Multiple-Output DC-DC Conversion for Low-Voltage Systems.” *IEEE Transactions on Very Large Scale Integration Systems*, vol. 8, pp. 252-263, June 2000.
- [14] Darling, Patrick. *Intel 22 nm 3D Tri-Gate Transistor Technology*. Technical Report, Intel Corporation, 2011.
- [15] Dennard, Robert H., et al. “Design of Ion-Implanted MOSFETs with Very Small Physical Dimensions.” *IEEE Journal of Solid State Circuits*, vol. 9, pp. 256-268, October 1974.
- [16] Deisch, Cecil. “Simple Switching Control Method Changes Power Converter into a Current Source.” *Power Electronics Specialists Conference*. vol. 1, 1978.
- [17] DiBene II, J. Ted, et al. “A 400 A Fully-Integrated Silicon Voltage Regulator with In-Die Magnetically Coupled Embedded Inductors.” *IEEE Applied Power Electronics Conference, Special Sessions on On-Die Voltage Regulators*. 2010.

- [18] Dickson, John F. "On-Chip High-Voltage Generation in MNOS Integrated Circuits Using an Improved Voltage Multiplier Technique." *IEEE Journal of Solid State Circuits*, vol. 11, pp. 374-378, June 1976.
- [19] Favrat, P., P. Deval, and M. J. Declercq. "A High Efficiency CMOS Voltage Doubler." *IEEE Journal of Solid State Circuits*, vol. 33, pp. 410-416, March 1998.
- [20] Fletcher, Jay and Steven Meyers. Programmable Bandgap Voltage Reference. USA Patent Pending. 2011.
- [21] Gardner, Donald S., et al. "Review of On-Chip Inductor Structures with Magnetic Films." *IEEE Transactions on Magnetics*. vol. 45, pp. 4760-4766, October 2009.
- [22] Greinacher, Heinrich. "Erzeugung Einer Gleichspannung Vom Vielfachen Betrage Einer Wechselfspannung Ohne Transformator." *Bulletin des Schweizer Elektotechnischer*, ver. 11, pp. 59-60, 1920.
- [23] Hales, Thomas C. "Cannonballs and Honeycombs." *Notices of the American Mathematical Society*. pp. 440-449, 2000.
- [24] Hilbiber, D. "A New Semiconductor Voltage Standard." *IEEE International Solid State Circuits Conference Digest of Technical Papers*. pp. 32-33, 1964.
- [25] Horowitz, Mark, Elad Alon, Dinesh Patel, Samuel Naffziger, Rajesh Kumar, and Kerry Bernstein. "Scaling, Power, and the Future of CMOS." *IEEE Electron Devices Meeting Technical Digest. IEEE International*. pp. 7-15, December 2005.
- [26] Howard, J., et al. "A 48-Core IA-32 Processor in 45 nm CMOS Using On-Die Message Passing and DVFS for Performance and Power Scaling." *IEEE Journal of Solid State Circuits*. vol. 46, pp. 173-18, 2011.
- [27] Ingerly, Doug, et al. "Low-k Interconnect Stack with Metal-Insulator-Metal Capacitors for 22nm High Volume Manufacturing." *IEEE International Interconnect Technology Conference (IITC)*. pp. 1-3, June 2012.

- [28] Intel Corporation. *Voltage Regulator-Down (VRD) 11.1: Processor Power Delivery Design Guidelines*, Document Number 322172-001, September 2009.
- [29] Johns, David, and Ken Martin. *Analog Integrated Circuit Design*. Wiley & Sons Inc, 1997.
- [30] Kurd, N. A., et al. "Westmere: A Family of 32 nm IA Processors." *IEEE International Solid States Circuits Conference Digest of Technical Papers*. pp. 96-97, 2011.
- [31] Kurson, Volkan, Siva Narendra, Vivek De, and Eby Friedman. "High Input Voltage Step-Down DC-DC Converters for Integration in a Low Voltage CMOS Process." *5th Int. Symp. on Quality Electronic Design*. pp. 517-521, 2004.
- [32] Le, Hanh-Phuc, Michael Seeman, Seth Sanders, Visvesh Sathe, Samuel Naffziger, and Elad Alon. "A 32 nm Fully Integrated Reconfigurable Switched Capacitor DC-DC Converter Delivering 0.55 W/mm² at 81% Efficiency." *IEEE International Solid State Circuits Conference Digest of Technical Papers*. pp. 210-211, 2010.
- [33] Le, Hanh-Phuc, Seth Sanders, and Elad Alon. "Design Techniques for Fully Integrated Switched-Capacitor DC-DC Converters." *IEEE Journal of Solid State Circuits*, vol. 46, pp. 2120-2131, July 2011.
- [34] Lee, Hoi, and Philip K. T. Mok. "An SC Voltage Doubler with Pseudo Continuous Output Regulation Using a Three Stage Switchable Op-Amp." *IEEE Journal of Solid State Circuits*, vol. 42, pp. 1216-1229, June 2007.
- [35] Ma, Dongsheng, and Feng Lo. "Robust Multiple Phase Switched Capacitor DC-DC Power Converter with Digital Interleaving Regulation Scheme." *IEEE Transactions on VLSI Systems*, vol. 16, pp. 611-619, June 2008.

- [36] Makowski, Marek S., and Dragan Maksimovic. "Performance Limits of Switched Capacitor DC-DC Converters." *IEEE Power Electronics Special Conference*. vol. 2, pp. 1215-1221, 1995.
- [37] Maksimovic, Dragan, and S. Dhar. "Switched-Capacitor DC-DC Converters for Low-Power On-Chip Applications." *IEEE Power Electronics Specialists Conference*. vol. 1, pp. 54-59, 1999.
- [38] Morrow, Patrick R., et al. "Design and Fabrication of On-Chip Coupled Inductors Integrated with Magnetic Material for Voltage Regulators." *IEEE Transactions on Magnetics*, vol. 47, pp. 1678-1686, February 2011.
- [39] Murata Manufacturing Company, Limited. "LQW18AN22NJ80 Inductor Datasheet." Product Datasheet. www.murata.com. April 2013.
- [40] Nagel, Laurence, and D. O. Pederson. *SPICE (Simulation Program with Integrated Circuit Emphasis)*. Technical Report, EECS Department, University of California Berkeley, UCB, 1973.
- [41] Nakagome, Yoshinobu, et al. "An Experimental 1.5-V 64-Mb DRAM." *IEEE Journal of Solid State Circuits*, vol. 26, pp. 465-472, April 1991.
- [42] Nose, Koichi and Takayasu Sakurai. "Optimization of V_{DD} and V_{TH} for Low-Power and High-Speed Applications." *Proceedings of the IEEE Asia and South Pacific Design Automation Conference*. pp. 469-474, June 2000.
- [43] Parry, John, Jeff Kotowski, and William McIntyre. Capacitor DC-DC Converter with PFM And Gain Hopping. USA Patent 6,055,168. April 25, 2000.
- [44] Patounakis, G., Y. W. Li, and K. L. Shepard. "A Fully Integrated On-Chip DC-DC Conversion and Power Management System." *IEEE Journal of Solid State Circuits*, vol. 39, pp. 443-451, March 2004.

- [45] Perrault, David J., and John Kassakian. "Distributed Interleaving of Paralleled Power Converters." *IEEE Transactions on Circuits and Systems I*, vol. 44, 728-734, August 1997.
- [46] Ramadass, Yogesh, and Anantha Chandrakasan. "Voltage Scalable Switched Capacitor DC-DC Converter for Ultra Low Power On-Chip Applications." *IEEE Power Electronics Specialists Conference*. pp. 2353-2359, June 2007.
- [47] Ramadass, Yogesh, Ayman Fayed, Baher Haroun, and Anantha Chandrakasan. "A 0.16 mm² Completely On-Chip Switched-Capacitor DC-DC Converter Using Digital Capacitance Modulation For LDO Replacement in 45 nm CMOS." *IEEE International Solid State Circuits Conference Digest of Technical Papers*. pp. 208-209, February 2010.
- [48] Rao, Arun. "An Efficient Switched Capacitor Buck-Boost Voltage Regulator Using Delta-Sigma Control Loop." Master's Thesis, Oregon State University, 2002.
- [49] Rao, Arun, W. McIntyre, Un-Ku Moon, and Gabor Temes. "Noise-Shaping Techniques Applied to Switched Capacitor Voltage Regulators." *IEEE Journal of Solid State Circuits*, vol. 40, pp. 422-429, February 2005.
- [50] Rawall, Bharat and Chris Reynolds. "RF Thin film Passive Devices." *Electronic Component News*. Rockaway, NJ, October 2009.
- [51] Razavi, Behzad. *Design of Analog CMOS Integrated Circuits*. New York, NY: McGraw-Hill Higher Education, 2001.
- [52] Razavi, Behzad. *Principles of Data Conversion System Design*. New York, NY: Wiley IEEE Press, 1995.
- [53] Schrom, Gerhard, F. Faillet, and Jaehong Hahn. "A 60 MHz 50 W Fine-Grain Package Integrated VR Powering A CPU from 3.3 V." *IEEE Special Sessions on On-Die Voltage Regulators in Applied Power Electronics Conference* . 2010.

- [54] Seeman, Michael, and Seth Sanders. "Analysis and Optimization of Switched Capacitor DC-DC Converters." *IEEE Transactions on Power Electronics*, vol. 23, pp. 841-851, 2008.
- [55] Seki, Akihisa. "PCBs With Embedded Components Emerge for Capacitors." *Asia Electronics Industry*, pp. 24-25, March 2011.
- [56] Somasekhar, Dinesh, et al. "Multi-phase 1 GHz Voltage Doubler Charge Pump in 32 nm Logic Process." *IEEE Journal of Solid State Circuits*, vol. 45, pp. 751-758, April 2010.
- [57] Sturcken, Noah; Davies, R.; Cheng Cheng; Bailey, William E.; Shepard, K.L., "Design of coupled power inductors with crossed anisotropy magnetic core for integrated power conversion." *IEEE Applied Power Electronics Conference and Exposition (APEC)*, pp. 417-423, Feb. 2012
- [58] Tesla, Nikola. Apparatus for Producing Currents of High Frequency. USA Patent 583,953. June 8, 1897.
- [59] Traff, H. "Novel Approach to High Speed CMOS Current Comparators." *Electronic Letters*. vol. 28, pp. 310-312, January 1992.
- [60] Vishay Intertechnology Inc. "TJ5-HT: Toroid, High Current, High Temperature, Radial Leaded Inductor." Product Datasheet. www.vishay.com. March 2012.
- [61] Wu, Chi-Hao, and Chern-Lin Chen. "A Low-Ripple Charge Pump with Continuous Pumping Current Control." *IEEE Proceedings of 51st Midwest Symposium on Circuits and Systems*. pp. 722-725, August 2008.
- [62] Zhang, Michael. "Powering Intel® Pentium® 4 generation processors." *Electrical Performance of Electronic Packaging*. pp. 215-218, 2001.
- [63] Zhang, Xiwen, and Hoi Lee. "An 88% Power Efficiency Accuracy-Enhanced DC-DC Conversion System for Transcutaneous-Powered Cochlear Implants." *IEEE*

Proceedings of Biomedical Circuits and Systems Conference. pp. 126-129, November 2007.

- [64] Zhang, Xiwen, and Hoi Lee. “An Efficiency-Enhanced Auto-Reconfigurable 2x/3x SC Charge Pump for Transcutaneous Power Transmission.” *IEEE Journal of Solid State Circuits*, vol. 45, pp. 1906-1922, September 2010.